FRAGMENTS AND ACTIVITY OF REL PROTEIN IN *M. TUBERCULOSIS* AND OTHER USES THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims priority of Application Serial No. 60/420,129, filed 2002 October 22 which is hereby incorporated by reference in its entirety.

GOVERNMENT SUPPORT

[0002] This invention was made with U.S. Government support (NIH Grant No. AI43420) and the U.S. Government may therefore have certain rights in the invention.

FIELD OF THE INVENTION

[0003] This invention is directed to the identification of antimycobacterial targets and the inhibition thereof as a treatment for infectious diseases, and more particularly, to the identification and inhibition of Rel protein in mycobacterium as a treatment for infectious diseases. The invention is also directed to compositions comprising attenuated mycobacteria that can be used to induce immune responses in an individual.

BACKGROUND

[0004] Tuberculosis (TB) in all of its manifestations is the leading cause of death from a single infectious agent. Studies from two urban centers indicate that 30-40% of new cases are the result of recent infection rather than reactivation of old disease, and cases acquired by recent transmission accounted from almost two-third's of drug resistant TB.

[0005] Highly resistant strains of *M. tuberculosis* have been isolated from patients in the Philadelphia area at a rate that requires physicians to treat every new case of presumed TB with at least four drugs, i.e., to consider every new case as if it were caused by one of these resistant strains. Clearly, new approaches to the development of antituberculosis therapy are necessary. However, the difficulties of working with Mycobacterium tuberculosis has kept the field from developing apace with the advances in molecular biology and biotechnology. In particular, the

analysis of the regulation of DNA replication, traditionally a rich area for the discovery of new antimicrobial agents and one that would provide major new insights into the growth of Mycobacteria, has been slow to develop.

[0006] The complete biochemistry of mycobacteria is not fully understood which makes it difficult to combat and destroy harmful mycobacteria. Current attempts in treating the infectious diseases resulting from harmful mycobacteria (*e.g. M. tuberculosis*), rely on destroying conventional targets within the mycobacteria, such as attacking the cell wall of the mycobacteria. Such practices however although may be therapeutic, however, are physically tasking to patients and are becoming increasingly marginalized as new and resistant strains of harmful mycobacteria are being discovered. Moreover, current practices do little to nothing to address the dormancy of harmful mycobacteria. Some harmful mycobacteria can exist, *in vivo*, in a dormant state waiting for more favorable environmental conditions to resurface and cause havoc.

[0007] There is a need to identify and inhibit new antimycobacterial targets that will yield better results than currently inhibited or treated antimycobacterial targets. Thus, there is a need to identify and inhibit new antimycobacterial targets that will be more effective in combating and destroying mycobacteria that persist *in vivo* in an active state and in a dormant state.

SUMMARY OF THE INVENTION

[0008] The present invention provides methods of inhibiting mycobacterial growth in a patient comprising administering an amount of a composition comprising an inhibitor of Rel_{Mtb}.

[0009] In some embodiments, the present invention provides recombinant vaccines comprising a nucleotide sequence that encodes *M. tuberculosis* immunogen operably linked to a regulatory elements, wherein said immunogen comprises Rel_{Mtb} or a fragment thereof.

[0010] In some embodiments, the present invention provides methods of inducing an immune response in a patient against a pathogen comprising administering to said patient a recombinant vaccine.

[0011] In some embodiments, the present invention provides methods of inducing an immune response in an animal comprising administering an immunogenic composition comprising an attenuated *M. tuberculosis*.

[0012] In some embodiments, the present invention provides methods of inducing an immune response in a patient comprising administering to said patient an immunogenic composition comprising a polypeptide, said polypeptide comprising a Rel_{Mtb} protein or fragment thereof.

[0013] In some embodiments, the present invention provides methods of modulating transcription in *M. tuberculosis* comprising administering to said *M. tuberculosis* a composition comprising a modulator of Rel_{Mtb}.

[0014] In some embodiments, the present invention provides isolated polynucleotides comprising a nucleotide sequence encoding a polypeptide comprising a fragment of a Rel_{Mtb} protein.

[0015] In some embodiments, the present invention provides expression vectors comprising a nucleic acid molecule that encodes for Rel_{Mtb} or a fragment thereof.

[0016] In some embodiments, the present invention provides host cells.

[0017] In some embodiments, the present invention provides antisense oligonucleotides comprising at least 80% sequence homology to the complement of a nucleic acid molecule encoding a *M. tuberculosis* Rel_{Mtb} polypeptide (SEQ ID NO:7, 8, 9, 10, 11, 12, or 13), wherein said antisense oligonucleotide specifically hybridizes to the nucleic acid molecule and inhibits *M. tuberculosis* Rel_{Mtb} mRNA levels by at least 50% in *M. tuberculosis*.

[0018] In some embodiments, the present invention provides compositions comprising a nucleic acid molecule of the present invention and a pharmaceutically acceptable carrier or diluent.

[0019] In some embodiments, the present invention provides compositions comprising a recombinant expression vector of the present invention and a pharmaceutically acceptable carrier or diluent.

[0020] In some embodiments, the present invention provides methods of producing a polypeptide that comprises a fragment of SEQ ID NO:7, said method comprising the steps of:

a) introducing a recombinant expression vector of the present invention into a compatible host cell;

b) growing said host cell under conditions for expression of said polypeptide; and

c) recovering said polypeptide.

[0021] In some embodiments, the present invention provides isolated polypeptide fragments of Rel_{Mtb} encoded by a nucleic acid molecule comprising an amino acid sequence with selected from the group consisting of SEQ ID NOs 14, 15, 16, 17, 18, and 19.

[0022] An isolated polypeptide fragment of Rel_{Mtb}, wherein said fragment has hydrolytic and/or synthetic activity, said hydrolytic and/or synthetic activity is not increased or decreased by more than 10% by the Rel_{Mtb} activating complex (RAC).

[0023] In some embodiments, the present invention provides compositions comprising a polypeptide of the present invention and an acceptable carrier or diluent.

[0024] In some embodiments, the present invention provides isolated antibodies which binds to an epitope on a polypeptide of the present invention.

[0025] In some embodiments, the present invention provides methods of identifying modulators of Rel_{Mtb} activity comprising:

a) contacting Rel_{Mtb} with a potential modulator; and

b) measuring the activity of Rel_{Mtb}

wherein if an activity of Rel_{Mtb} is inhibited then the modulator is an inhibitor of Rel_{Mtb} activity, and if an activity of Rel_{Mtb} is increased the modulator is an activator of Rel_{Mtb} activity.

[0026] In some embodiments, the present invention provides methods of protecting a patient from a *M. tuberculosis* infection comprising administering to said patient an amount of a composition comprising a RelMtb modulator effective to protect the animal from *M. tuberculosis* infection.

[0027] In some embodiments, the present invention provides methods of modulating growth of a pathogen comprising administering to said pathogen an amount of a composition comprising a modulator effective to inhibit growth of said pathogen.

[0028] In some embodiments, the present invention provides methods of inhibiting dormancy in *M. tuberculosis* comprising administering to said *M. tuberculosis* a composition comprising a modulator of Rel_{Mtb}.

[0029] In some embodiments, the present invention provides attenuated *M. tuberculosis* comprising an inactivated Rel_{Mtb} gene, wherein said inactivated Rel_{Mtb} gene encodes for a Rel_{Mtb} protein with diminished synthetic activity, hydrolytic activity, or both.

BRIEF DESCRIPTION OF THE DRAWINGS:

[0030] Figure 1: Summary of Rel_{Mtb} truncated proteins. Full-length Rel_{Mtb} protein is at the top followed by the different truncated proteins. Amino acid numbers are at the beginning and end of each fragment and corresponding activity is listed below. 87-181 overlapping site is noted in the full-length Rel_{Mtb}.

[0031] Figure 2: Dependence of the initial velocity (vo) of Rel_{Mtb} fragment 1-394 transferase reaction on [Mg2+] and [Mn2+]. (A) vo as a function of [Mg2+] and a fixed substrate concentration (1mM ATP + 1mM GTP) at a fixed pH (8.0). (B) vo as a function of [Mn2+] and a fixed substrate concentration (1mM ATP + 1mM GTP) at a fixed pH (8.0).

[0032] Figure 3: Dependence of the initial velocity (vo) of hydrolysis reaction on [Mn2+] for Rel_{Mtb} fragments. (A) 1-394 fragment vo as a function of [Mn2+] and 0.3 mM (●), 0.5 mM (■), and 0.8 mM (▲) pppGpp at a fixed pH (8.0). (B) 1-203 fragment vo as a function of [Mn2+] and 0.3 mM (●), 0.5 mM (■) pppGpp at a fixed pH (8.0). (C) 1-181 fragment vo as a function of [Mn2+] and 0.5 mM pppGpp at a fixed pH (8.0).

[0033] Figure 4: HPLC elution (200 mM NaCl, 50 mM HEPES, pH 8.2) monitored at A280 for (A) full-length Rel_{Mtb} (50 μg). Retention time (16.72 s) corresponds to 240 kDa protein. (B) for 1-394 fragment (96 μg). Retention times (18.1 s and 21.15 s) correspond to 141 kDa (peak 1) and 47 kDa (peak 2). (C) for 1-203 fragment (258 μg). Retention time (22.28 s) corresponds to 31.7 kDa protein. (D) for 1-181 fragment (87.5 μg). Retention time (22.64 s) corresponds to 27.8 kDa protein.

[0034] Figure 5: Summary of truncated Rel_{Mtb} proteins with labeled domains. Figure 5A is full-length Rel_{Mtb} and 5B-G are truncated proteins. Second column lists corresponding activity and third column lists important features of protein followed by appropriate reference.

[0035] Figure 6: Shared similarities in the amino acid starvation response pathways of prokaryotic Rel_{Mtb} and eukaryotic GCN2.

DETAILED DESCRIPTION

[0036] An Mtb (*M. tuberculosis*) gene, which we designate Rel_{Mtb}, encodes a protein of 738 amino acid residues and belongs to the relA/spoT family of genes that mediates the stringent response (5). A knockout strain of Rel_{Mtb} (ΔRelMtb) demonstrated that the enzyme is responsible for the intracellular regulation of (p)ppGpp, the effector of the stringent response, and the consequent ability of Mtb to survive long-term starvation in culture (4).

[0037] Rel_{Mtb} is a member of the superfamily of single-gene encoded bifunctional enzymes that catalyzes opposing reactions (6). During amino acid starvation, the Rel_{Mtb} protein catalyzes the transfer of the 5'-β,γ pyrophosphate group from ATP to the 3'-OH of GTP: ATP + GTP ↔ AMP + pppGpp. When amino acid levels return to normal, the stringent response is reversed by Rel_{Mtb} catalyzing the hydrolysis of the pyrophosphate group (PPi) from the 3'-OH of both pppGpp and

ppGpp yielding GTP or GDP: (p)ppGpp → GTP (GDP) + PPi (4,6). Rel_{Mtb} is differentially regulated depending on the aminoacylation state of a tRNA·ribosome·mRNA complex (Rel_{Mtb} activating complex or RAC) (6). Without efficient Rel_{Mtb} regulation there would be a futile cycling of ATP and GTP, resulting in a drastic decrease of high-energy compounds necessary for cellular function.

[0038] The dual-function Rel_{Mtb} protein from Mycobacterium tuberculosis catalyzes both the synthesis and hydrolysis of (p)ppGpp, the effector of the stringent response. Previous work (Avarbock, D., Avarbock, A. and Rubin, H. (2000) Biochemistry 39, 11640) presented evidence that the full-length 738 amino acid (82 kDa) Rel_{Mtb} protein may catalyze its two opposing reactions at two distinct active sites, although the sites were not identified. In the present invention, we have purified, characterized, and identified fragments of Rel_{Mtb} protein as well as the full length protein and identified that Rel_{Mtb} catalyzes its two opposing reactions at distinct sites. A fragment containing amino acids 87-394 (35 kDa fragment) has only (p)ppGpp synthesis activity and a fragment containing amino acids 1-181 (20 kDa fragment) has only (p)ppGpp hydrolysis activity. We also purified a fragment containing amino acids 1-394 (45 kDa fragment) that possesses both synthesis and hydrolysis activities. Unlike wild type Rel_{Mtb}, the synthesis activity of the fragments is not enhanced by the previously described Rel_{Mtb} activating complex (RAC), and the hydrolysis activity of the fragments is not inhibited by this complex. The basal hydrolysis kcat/Km of the fragments is decreased approximately 55 fold compared to the basal hydrolysis activity of the wild type Rel_{Mtb}, whereas kcat/Km for synthesis only decreased approximately 2 fold. In addition, wild type Rel_{Mtb} exclusively forms trimers and removal of the C-terminus results in the isolation of monomers. Therefore, Rel_{Mtb} catalyzes two opposing reactions at distinct active sites, and the C-terminus is involved in multimerization and regulation of both synthesis and hydrolysis.

[0039] Various definitions are made throughout this document. Most words have the meaning that would be attributed to those words by one skilled in the art. Words specifically defined either below or elsewhere in this document have the meaning provided in the context of the present invention as a whole and as are typically understood by those skilled in the art.

[0040] Accession numbers disclosed herein are each incorporated by reference in their entirety. For example, *M. tuberculosis* accession numbers begin with "Rv" (*e.g.* Rv2416c) and are, therefore, in some instances referred to as the "Rv number."

[0041] As used herein, "modulation" means either an increase (stimulation) or a decrease (inhibition) in the expression of a gene or in the amount, quality, or effect of a particular activity

or protein. In some embodiments of the present invention, inhibition is the form of modulation of gene expression.

[0042] As used herein, the term "inhibit" refers to a reduction or decrease in an activity, quality or quantity, compared to a baseline. For example, in the context of the present invention, inhibition of viral replication refers to a decrease in viral replication as compared to baseline. In some embodiments there is a reduction of about 30%, about 40%, about 50%, about 60%, about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, about 99%, and about 100%. Those of ordinary skill in the art can readily determine whether or not viral replication has been inhibited and to what extent.

[0043] As used herein, the terms "modulator," "inhibitor," "stimulator," or "activator" refers to compounds that can either modulate, inhibit, stimulate, or activate a function or activity. In some embodiments, the compounds are antibodies, peptides, polypeptides, small molecular weight compounds, antisense compounds, or RNAi compounds. In some embodiments, the antisense compound is an antisense oligonucleotide.

[0044] As used herein, the term "selectively inhibit" refers to selective inhibition of *M. tuberculosis* relative to other biological targets. For example, in some embodiments, the inhibitors of the present invention selectively inhibit *M. tuberculosis* but does not inhibit the dopamine pathway. In some embodiments, the affinity of the inhibitors of the present invention is at least 10, at least 10^2, or at least 10^3-fold higher for Rel_{Mtb} over another target. In some embodiments, the selectivity is determined by comparing IC_{50}, EC_{50}, MIC, or the like, between *in vivo* or *in vitro* models of *M. tuberculosis* cell growth, electron transport, cell respiration, cell replication, and *in vivo* or *in vitro* models of the human dopamine pathway.

[0045] As used herein, the term "about" refers to +/- 20%, +/- 15%, +/- 10%, or +/- 5% of a given value.

[0046] The articles "a" and "an" are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

[0047] As used herein, the term "prophylactically effective amount" is meant an amount of an inhibitor of the present invention effective to yield the desired prophylactic response. The specific prophylactically effective amount will, obviously, vary with such factors as the physical condition of the patient, the type of mammal or animal being treated, the duration of the treatment, the nature of concurrent therapy (if any), and the specific formulations employed and the structure of the inhibitor or its derivatives.

[0048] As used herein "mycobacteria" refers to any bacteria that falls within the mycobacterial genus. In some embodiments the mycobacteria is *M. tuberculosis* or *M. smegmatis.*

[0049] As used herein "combination therapy" means that the individual in need of treatment is given another medicament for the disease in conjunction with the inhibitors of the present invention. This combination therapy can be sequential therapy where the individual is treated first with one or more drugs and then the other, or two or more drugs are given simultaneously.

[0050] As used herein, the term "an individual suspected of having been exposed to one or more pathogens" refers to an individual who has not been diagnosed as being positive for one or more pathogens but who could possibly have been exposed to one or more pathogens due to a recent high risk activity or activity that likely put them in contact with the pathogens. For example, an individual suspected of having been exposed to *M. tuberculosis* refers to an individual that has been in close proximity to an individual infected with *M. tuberculosis* or having had contact or exposure to samples comprising *M. tuberculosis.* Examples of such samples include, without limitation, laboratory or research samples or samples of blood, semen, bodily secretions, and the like from patients. The individual from which the pathogen originated may or may not have been tested for the presence and/or absence of the pathogen. The term "an individual suspected of having been exposed to one or more pathogens" also includes individuals who have been diagnosed as being positive for one pathogen but are also infected with at least one further pathogen.

[0051] As used herein, the term "sample" refers to biological material from a patient. The sample assayed by the present invention is not limited to any particular type. Samples include, as non-limiting examples, single cells, multiple cells, tissues, tumors, biological fluids, biological molecules, or supernatants or extracts of any of the foregoing. Examples include tissue removed for biopsy, tissue removed during resection, blood, urine, saliva, lymph tissue, lymph fluid, cerebrospinal fluid, mucous, and stool samples. The sample used will vary based on the assay format, the detection method and the nature of the tumors, tissues, cells or extracts to be assayed. Methods for preparing samples are well known in the art and can be readily adapted in order to obtain a sample that is compatible with the method utilized.

[0052] The terms "subject" or "patient" as used herein include any mammalian species. In some embodiments, the methods of the present invention are contemplated for the treatment of infectious diseases in mammals such as humans, as well as those mammals of importance due to being endangered, of economical importance and/or social importance to humans.

[0053] "Rel_{Mtb} activity" as used herein refers to hydrolysis or synthesis of (p)ppGpp. "Rel_{Mtb} activity" can also refer to the effect on transcription of other genes. "Synthesis activity" can also be referred to as "transferase activity" or "synthetic activity."

[0054] "Diminished activity" refers to activity that is less than what is observed with wild-type or full length Rel_{Mtb} protein. In some embodiments the activity is reduced about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 100%. In some embodiments, the diminished activity refers to either the hydrolysis activity or the synthesis activity. In some embodiments "diminished activity" refers to both the hydrolysis activity and the synthesis activity. "Diminished activity" can also refer to the transcription levels of specific genes that are affected by the activity of Rel_{Mtb}. The "diminished activity" can also refer to fragments of Rel_{Mtb} protein.

[0055] "Stimulated activity" refers to activity that is more than what is observed with wild-type or full length Rel_{Mtb} protein. In some embodiments the activity is increased about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 100%. In some embodiments, the stimulated activity refers to either the hydrolysis activity or the synthesis activity. In some embodiments "stimulated activity" refers to both the hydrolysis activity and the synthesis activity. "Stimulated activity" can also refer to the transcription levels of specific genes that are affected by the activity of Rel_{Mtb}. The "stimulated activity" can also refer to fragments of Rel_{Mtb} protein.

[0056] "Synthesized nucleotides" as used herein and understood in the art, refers to polynucleotides produced by purely chemical, as opposed to enzymatic, methods. "Wholly" synthesized DNA sequences are therefore produced entirely by chemical means, and "partially" synthesized DNAs embrace those wherein only portions of the resulting DNA were produced by chemical means.

[0057] As used herein, the term "region" is meant a physically contiguous portion of the primary structure of a biomolecule. In the case of proteins, a region is defined by a contiguous portion of the amino acid sequence of that protein. In the case of polynucleotides, a region is defined by a contiguous portion of the nucleotide sequence of that polynucleotide. Examples of polynucleotide regions include without limitation, the 5'UTR, the start codon region, an intron/exon region, the stop codon region, and the 3'UTR. For example, the term "start codon region" refers to a portion of a polynucleotide that encompasses from about 25 to about 50 contiguous nucleotides in either direction (i.e., 5' or 3') from a start codon. Similarly, the term

"stop codon region" refers to a portion of a polynucleotide that encompasses from about 25 to about 50 contiguous nucleotides in either direction from a stop codon.

[0058] The term "domain" as used herein defined as referring to a structural part of a biomolecule that contributes to a known or suspected function of the biomolecule. Domains may be co-extensive with regions or portions thereof; domains may also incorporate a portion of a biomolecule that is distinct from a particular region, in addition to all or part of that region . Examples of Rel_{Mtb} protein domains include, but are not limited to, hydrolysis domain, synthesis domain, and ribosomal regulation domain.

[0059] As used herein, the term "activity" refers to a variety of measurable indicia suggesting or revealing binding, either direct or indirect; affecting a response, i.e. having a measurable affect in response to some exposure or stimulus, including, for example, the affinity of a compound for directly binding a polypeptide or polynucleotide of the invention, or, for example, measurement of amounts of upstream or downstream proteins or other similar Function: S after some stimulus or event.

[0060] As used herein, the term "antibody" is meant to refer to complete, intact antibodies, and Fab, Fab', F(ab)2, and other fragments thereof. Complete, intact antibodies include monoclonal antibodies such as murine monoclonal antibodies, chimeric antibodies and humanized antibodies.

[0061] As used herein, the term "binding" means the physical or chemical interaction between two proteins or compounds or associated proteins or compounds or combinations thereof. Binding includes ionic, non-ionic, Hydrogen bonds, Van der Waals, hydrophobic interactions, etc. The physical interaction, the binding, can be either direct or indirect, indirect being through or due to the effects of another protein or compound. Direct binding refers to interactions that do not take place through or due to the effect of another protein or compound but instead are without other substantial chemical intermediates. Binding may be detected in many different manners. As a non-limiting example, the physical binding interaction between a Rel_{Mtb} of the invention and a compound can be detected using a labeled compound. Methods of detecting binding are well known to those of skill in the art.

[0062] As used herein, the term "compound" means any identifiable chemical or molecule, including, but not limited to, small molecule, peptide, protein, sugar, nucleotide, or nucleic acid, and such compound can be natural or synthetic.

[0063] As used herein, the term "complementary" refers to Watson-Crick basepairing between nucleotide units of a nucleic acid molecule.

[0064] As used herein, the term "contacting" means bringing together, either directly or indirectly, a compound into physical proximity to a polypeptide or polynucleotide of the invention. The polypeptide or polynucleotide can be in any number of buffers, salts, solutions etc. Contacting includes, for example, placing the compound into a beaker, microtiter plate, cell culture flask, or a microarray, such as a gene chip, or the like, which contains the nucleic acid molecule, or polypeptide encoding the Rel_{Mtb} or fragment thereof.

[0065] As used herein, the phrase "homologous nucleotide sequence," or "homologous amino acid sequence," or variations thereof, refers to sequences characterized by a homology, at the nucleotide level or amino acid level, of at least the specified percentage. Homologous nucleotide sequences include those sequences coding for isoforms of proteins. Such isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. Homologous nucleotide sequences include nucleotide sequences encoding for a protein of a species other than *M. tuberculosis*, including, but not limited to, other bacteria and mammals. Homologous amino acid sequences include those amino acid sequences which contain conservative amino acid substitutions and which polypeptides have the same binding and/or activity. Percent homology can be determined by, for example, the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison WI), using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482-489, which is incorporated herein by reference in its entirety).

[0066] As used herein, the term "isolated" nucleic acid molecule refers to a nucleic acid molecule (DNA or RNA) that has been removed from its native environment. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules.

[0067] As used herein, the terms "modulates" or "modifies" means an increase or decrease in the amount, quality, or effect of a particular activity or protein.

[0068] As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues which has a sufficient number of bases to be used in a polymerase chain reaction (PCR). This short sequence is based on (or designed from) a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. In some embodiments, oligonucleotides comprise portions of a DNA or RNA sequence having at least about 10 nucleotides and as many as about 50

nucleotides, at least about 10 nucleotides and as many as about 40 nucleotides, at least about 10 nucleotides and as many as about 30 nucleotides. They are chemically synthesized and may be used as probes.

[0069] As used herein, the term "probe" refers to nucleic acid sequences of variable length, for example, between at least about 10 and as many as about 6,000 nucleotides, depending on use. They are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and much slower to hybridize than oligomers. They may be single- or double-stranded and carefully designed to have specificity in PCR, hybridization membrane-based, or ELISA-like technologies.

[0070] The term "preventing" refers to decreasing the probability that an organism contracts or develops an abnormal condition.

[0071] The term "treating" refers to having a therapeutic effect and at least partially alleviating or abrogating an abnormal condition in the organism.

[0072] The term "therapeutic effect" refers to the inhibition or activation factors causing or contributing to the abnormal condition. A therapeutic effect relieves to some extent one or more of the symptoms of the abnormal condition. In reference to the treatment of abnormal conditions, a therapeutic effect can refer to one or more of the following: (a) an increase in the proliferation, growth, and/or differentiation of cells; (b) inhibition (i.e., slowing or stopping) of cell death; (c) inhibition of degeneration; (d) relieving to some extent one or more of the symptoms associated with the abnormal condition; and (e) enhancing the function of the affected population of cells. Compounds demonstrating efficacy against abnormal conditions can be identified as described herein.

[0073] The term "abnormal condition" refers to a function in the cells or tissues of an organism that deviates from their normal Function: S in that organism. An abnormal condition can relate to cell proliferation, cell differentiation, cell signaling, or cell survival. An abnormal condition may also include obesity, diabetic complications such as retinal degeneration, and irregularities in glucose uptake and metabolism, and fatty acid uptake and metabolism.

[0074] Abnormal cell survival conditions may also relate to conditions in which programmed cell death (apoptosis) pathways are activated or abrogated. A number of protein kinases are associated with the apoptosis pathways. Aberrations in the function of any one of the protein kinases could lead to cell immortality or premature cell death.

[0075] The term "administering" relates to a method of incorporating a compound into cells or tissues of an organism. The abnormal condition can be prevented or treated when the cells or tissues of the organism exist within the organism or outside of the organism. Cells existing outside the organism can be maintained or grown in cell culture dishes. For cells harbored within the organism, many techniques exist in the art to administer compounds, including (but not limited to) oral, parenteral, dermal, injection, and aerosol applications. For cells outside of the organism, multiple techniques exist in the art to administer the compounds, including (but not limited to) cell microinjection techniques, transformation techniques and carrier techniques.

[0076] The abnormal condition can also be prevented or treated by administering a compound to a group of cells having an aberration in a signal transduction pathway to an organism. The effect of administering a compound on organism function can then be monitored. The organism in some embodiments is a mouse, rat, rabbit, guinea pig or goat, more preferably a monkey or ape, and most preferably a human.

[0077] As used herein, the phrase "stringent hybridization conditions" or "stringent conditions" refers to conditions under which a probe, primer, or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present in excess, at Tm, 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (*e.g.* 10 to 50 nucleotides) and at least about 60°C for longer probes, primers or oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

[0078] The amino acid sequences are presented in the amino to carboxy direction, from left to right. The amino and carboxy groups are not presented in the sequence. The nucleotide sequences are presented by single strand only, in the 5' to 3' direction, from left to right. Nucleotides and amino acids are represented in the manner recommended by the IUPAC-IUB Biochemical Nomenclature Commission or (for amino acids) by three letters code.

[0079] Based on the studies disclosed herein, the Rel protein plays a role in survival for *M. tuberculosis*. Thus, if the function of Rel can be modulated (*e.g.* activated or inhibited) one can treat or ameliorate animals that are infected with the bacteria that causes tuberculosis.

[0080] Therefore, the present invention provides methods of inhibiting mycobacterial growth. The methods comprise administering an effective amount of a composition comprising an modulator of Rel_{Mtb} to a cell or an animal. In some embodiments the modulator inhibits or stimulates the hydrolysis activity of the Rel_{Mtb} protein. In some embodiments, the modulator inhibits or stimulates the synthesis activity of the Rel_{Mtb} protein. In some embodiments, the modulator inhibits or stimulates the synthesis activity and the hydrolysis activity of the Rel_{Mtb} protein. In some embodiments, the mycobacteria is tuberculosis, although it is envisioned that the modulators can be used to inhibit the growth of other types of bacteria that cause infections such as staph and strep, which are common infectious causing agents. In some embodiments the modulator is co-administered with at least one additional antibiotic (*e.g.* Isoniazid and/or Rifampicin). It is well known to those of skill in the art how to administer a composition to an animal to treat a disease or disorder. In some embodiments the animal is a human, mouse, dog, cat, or horse. In some embodiments, the inhibitor is used to inhibit the growth of *S. aureus* , *E faecalis*, *H. influenzae*, *M catarrhalis*, *S. pneumoniae*, and *E. Coli*.

[0081] Administering an inhibitor of Rel_{Mtb} can also be used to ameliorate a mycobacterial infection. Since, the Rel_{Mtb} protein can play a role in pathogenesis of tuberculosis inhibiting the function of Rel_{Mtb} can lead to a reduction of the disease. In some embodiments, the inhibition of Rel_{Mtb} can make the bacteria more susceptible to an additional compound or antibiotic, which would inhibit the growth of the bacteria.

[0082] Measuring the growth of mycobacteria is well known to those of skill in the art and can be done by any appropriate means. In some embodiments, mycobacterial growth is monitored by oxygen consumption. Oxygen consumption can be monitored using the BACTEC™ system, which monitors oxygen consumption with fluorescence quenching. Other methods that are effective for measuring the growth of mycobacteria include, but are not limited, using a radioactive carbon source as a nutrient and measuring radioactive carbon dioxide output as an indicator of growth. In some embodiments, other methods to measure mycobacterial growth include, but are not limited to, plating TB on growth medium and visualizing colony growth, growing mycobacteria in liquid culture and using a spectrometer to read the optical density (O.D. @ 600nm) of bacilli in liquid medium, putting radioactive nucleotides in the growth medium (such as radioactive uracil) and measuring the radioactivity of the bacilli.

[0083] One aspect of tuberculosis is that the bacteria can go dormant and remain dormant for years or decades. There have even been reports of tuberculosis causing bacteria being isolated from mummies that are hundreds of years old that are able to be reactivated. Thus, inhibiting dormancy could enable the bacteria to be more thoroughly cleared from a patient. Therefore, the present invention also provides methods of inhibiting dormancy in *M. tuberculosis*. In some embodiments, the method comprises administering to the *M. tuberculosis* a composition comprising a modulator of Rel_{Mtb}. "Dormancy" refers to the state *M. tuberculosis* enters in response to a lack of nutrients, oxygen, or other unfavorable conditions. As discussed herein, the modulator can be an antibody, peptide, polypeptide, small molecular weight organic compound, antisense compound, or RNAi compound. In some embodiments, the modulator can inhibit or activate the synthetic activity, hydrolytic activity, or the transcriptional activity of Rel_{Mtb}. In some embodiments the dormant bacteria is present within a patient.

[0084] "An effective amount" can also refer to an amount that inhibits the growth of mycobacteria in an *in vitro* assay. In some embodiments, the effective amount inhibits the growth of mycobacteria in an *in vitro* assay by at least 40%, 50%, 60%, 70%, 80%, 90%, 95%, 95%, 96%, 97%, 98%, 99%, or 100%. An *in vitro* assay can also refer to an assay that measures the hydrolysis or synthesis activity of the protein as described in the Examples section. However, any assay can be used to measure catalytic activity. An effective amount can also refer to an amount that has a therapeutic effect.

[0085] The present invention also provides recombinant vaccines for the prevention of tuberculosis. In some embodiments, a recombinant vaccine comprises a nucleotide sequence that encodes a *M. tuberculosis* immunogen operably linked to regulatory elements. In some embodiments the immunogen comprises Rel_{Mtb} protein or a fragment thereof.

[0086] "Fragment" as used herein can also refer to fragment of a protein that is also functional. For example, a functional fragment of Rel_{Mtb} can have hydrolysis activity, synthesis activity, or the combination of both. A fragment can also have no activity and can be referred to as a "non-functional" fragment. In some embodiments, the recombinant vaccine encodes an immunogen comprising a functional fragment of Rel_{Mtb} or a non-functional fragment of Rel_{Mtb}. In some embodiments, the nucleotide sequence encodes for a polypeptide comprising a sequence of SEQ ID NOs., 7, 8, 9, 10, 11, 12, or 13. In some embodiments, the recombinant vaccine comprises a nucleotide sequence comprising a sequence of SEQ ID NOs. 14, 15, 16, 17, 18, 19, or 20. A "fragment" of Rel_{Mtb} protein is less than the complete contiguous amino acid sequence of Rel_{Mtb}. For example, if a polypeptide comprises a fragment of Rel_{Mtb} it does not have the complete

contiguous amino acid sequence of Rel_{Mtb}. In some embodiments, the polypeptide comprising a fragment of Rel_{Mtb} can also comprise regions of another protein that is not Rel_{Mtb}. A "fragment of Rel_{Mtb}" as used herein can also refer to a fragment that has hydrolytic or synthetic activity, but the activity is not increased or decreased by the Rel_{Mtb} activating complex (RAC). In some embodiments, the activity is not increased or decreased by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%.

[0087] In some embodiments, the recombinant vaccine is a recombinant vaccinia vaccine.

[0088] The present invention also provides methods for inducing an immune response in an animal (*e.g.* human, mouse, dog, cat, monkey, or horse) against a pathogen (*e.g. M. tuberculosis* or *M. smegmatis*). comprising administering to the animal a recombinant vaccine.

[0089] The present invention provides methods of inducing an immune response in an animal comprising administering an immunogenic composition comprising an attenuated mycobacteria tuberculosis.

[0090] As used herein, "attenuated mycobacteria tuberculosis" refers to a live bacteria that is less pathogenic then its normal wild-type form. A bacteria can be attenuated by various means known to those of skill in the art. In some embodiments, the attenuated mycobacteria comprises an inactivated Rel_{Mtb} gene.

[0091] An "inactivated Rel_{Mtb} gene" refers to a gene that is either not expressed or whose normal expression is reduced compared to the non-attenuated bacteria. In some embodiments the inactivated gene's expression is about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 100% less than the non-attenuated bacteria. Expression can be measured using any means including, but not limited to, Western Blot, Southern Blot, Northern Blot, RT-PCR, and the like.

[0092] In some embodiments the attenuated bacteria comprises a mutated Rel_{Mtb} gene. A mutated Rel_{Mtb} gene can encode for a fragment. In some embodiments the mutation (*e.g.* insertion, deletion, frameshift, and the like) can result gene being having reduced expression as compared to the non-attenuated bacteria. In some embodiments, the protein that is encoded for by the mutated Rel_{Mtb} gene has diminished activity. In some embodiments, the hydrolysis activity is diminished. In some embodiments, the synthesis activity is diminished. In some embodiments the hydrolysis activity and the synthesis activity is diminished. In some embodiments the mutated Rel_{Mtb} gene encodes for a polypeptide having the sequence of SEQ ID NOs. 14, 15, 16, 17, 18, or 19. In some embodiments, the mutated Rel_{Mtb} gene encodes for a polypeptide that has stimulated activity (*e.g* hydrolysis and/or synthesis) as compared to the non-attenuated (*e.g.* wild-type) bacteria. In some embodiments, the attenuated bacteria has the Rel

gene deleted. In addition to the Rel gene, additional genes can be deleted. In some embodiments, the attenuated bacterial in addition to the Rel gene being deleted or inactivated, at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more gene are deleted or inactivated. In some embodiments, genes that are involved in electron transport are deleted and/or inactivated.

[0093] The present invention also provides methods of inducing an immune response in an animal comprising administering an immunogenic composition comprising a polypeptide, wherein the polypeptide comprises the Rel_{Mtb} protein or fragments thereof. In some embodiments, the fragments are functional and/or non-functional. In some embodiments more than one polypeptide is administered to the animal for an immune response. The fragments that can be administered can comprise a polypeptide having the sequence of SEQ ID NOs: 7, 8, 9, 10, 11, 12, and 13. In some embodiments, the polypeptides have diminished and/or stimulated Rel_{Mtb} activity.

[0094] The immunogenic compositions of the present invention can be administered by any method. In some embodiments, the immunogenic composition can be administered more than once. In some embodiments it is administered more than once over the span of one day, one week, one month, two months, three months, six months, or one year. The number of times that the composition is administered and when it is administered is routine for one of skill in the art. The immunogenic composition(s) can also be administered with at least one adjuvant to facilitate the generation of an immune response or to increase the immune response.

[0095] Inducing an immune response in an animal or a patient can also be used to induce a passive immune response and the antibodies that are generated can be collected from the animal or a patient. The antibodies can be used, *inter alia,* as research and diagnostic tools.

[0096] As discussed herein and below, an attenuated mycobacteria with a deleted Rel gene has an effect on the expression of numerous genes in bacteria. Therefore, the present invention provides methods of modulating the transcription in mycobacteria comprising administering a composition comprising a modulator of Rel. In some embodiments the modulator modulates the activity of the Rel protein. In some embodiments, the modulator is an inhibitor, activator, or stimulator.

[0097] As used herein "mycobacteria" refers to any bacteria that falls within the mycobacterial genus. In some embodiments the mycobacteria is *M. tuberculosis* or M. smegmatis.

[0098] As used herein, the terms "modulator," "inhibitor," "stimulator," or "activator" refers to compounds that can either modulate, inhibit, stimulate, or activate a function or activity. In some embodiments, the compounds are antibodies, peptides, polypeptides, small molecular

weight compounds, antisense compounds, or RNAi compounds. In some embodiments, the antisense compound is an antisense oligonucleotide.

[0099] In some embodiments an inhibitor inhibits an activity or function at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 100%. In some embodiments a stimulator increase an activity or function at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 100%. In some embodiments an activator activates an activity or function at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 100%. In some embodiments, a modulator modifies the activity or function either by increasing or decreasing the activity or function by at least In some embodiments a stimulator increase an activity or function at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 100%.

[00100] The present invention also provides methods of protecting an animal from mycobacterial infection comprising administering to the animal an amount of a composition comprising a Rel protein modulator effective to protect the animal from the infection.

[00101] The present invention further provides methods of modulating the growth of a pathogen (*e.g.* mycobacteria) comprising administering to the pathogen an amount of a composition comprising a modulator effective to inhibit the growth of the pathogen. In some embodiments the modulator is a Rel protein or Rel gene modulator. Examples of a Rel gene modulator include, but are not limited to antisense and RNAi compounds.

Polynucleotides

[00102] The present invention provides purified and isolated polynucleotides (*e.g.*, DNA sequences and RNA transcripts, both sense and complementary antisense strands, both single- and double-stranded) that encode Rel_{Mtb}. The gene described herein is referred to as Rel_{Mtb}.

[00103] The invention provides purified and isolated polynucleotides (*e.g.*, cDNA, genomic DNA, synthetic DNA, RNA, or combinations thereof, whether single- or double-stranded) that comprise a nucleotide sequence encoding the amino acid sequence of the polypeptides of the invention. Such polynucleotides are useful for recombinantly expressing the protein and also for detecting expression of the protein in cells (*e.g.*, using Northern hybridization and in situ hybridization assays). Such polynucleotides also are useful in the design of antisense and other molecules for the suppression of the expression of Rel_{Mtb} in a cultured cell, a tissue, or an animal; for therapeutic purposes; or to provide a model for diseases or conditions characterized by Rel_{Mtb} expression. Specifically excluded from the definition of polynucleotides of the invention are entire isolated, non-recombinant native chromosomes of host cells. In some

embodiments the polynucleotide has a sequence of SEQ ID NOs: 14, 15, 16, 17, 18, 19 or 20, which correspond to naturally occurring Rel_{Mtb} sequences. It will be appreciated that numerous other polynucleotide sequences exist that also encode Rel_{Mtb} protein having a sequence of SEQ ID NOs: 7, 8, 9, 10, 11, 12, or 13, due to the well-known degeneracy of the universal genetic code.

[00104] The invention also provides a purified and isolated polynucleotide comprising a nucleotide sequence that encodes a polypeptide, wherein the polynucleotide hybridizes to a polynucleotide having the sequence set forth in sequences of SEQ ID NOs: 14, 15, 16, 17, 18, 19 or 20, or the non-coding strand complementary thereto, under the following hybridization conditions:

[00105] (a) hybridization for 16 hours at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% dextran sulfate; and

[00106] (b) washing 2 times for 30 minutes each at 60°C in a wash solution comprising 0.1% SSC, 1% SDS. Polynucleotides that encode a human allelic variant are highly preferred.

[00107] The present invention relates to molecules which comprise the gene sequences that encode the Rel_{Mtb}; constructs and recombinant host cells incorporating the gene sequences; the novel Rel_{Mtb} polypeptides encoded by the gene sequences; antibodies to the polypeptides and homologs; kits employing the polynucleotides and polypeptides, and methods of making and using all of the foregoing. In addition, the present invention relates to homologs of the gene sequences and of the polypeptides and methods of making and using the same.

[00108] The invention also comprehends cDNA that is obtained through reverse transcription of an RNA polynucleotide encoding Rel_{Mtb} (conventionally followed by second strand synthesis of a complementary strand to provide a double-stranded DNA).

[00109] In some embodiments DNA sequences encoding Rel_{Mtb} polypeptides are selected from SEQ ID NOs: 14, 15, 16, 17, 18, 19 or 20. In some embodiments DNA of the invention comprises a double stranded molecule along with the complementary molecule (the "non-coding strand" or "complement") having a sequence unambiguously deducible from the coding strand according to Watson-Crick base-pairing rules for DNA. In some embodiments are other polynucleotides encoding the Rel_{Mtb} polypeptide of SEQ ID NO:7-13, which differ in sequence from the polynucleotides of SEQ ID NOs: 14, 15, 16, 17, 18, 19 or 20, by virtue of the well-known degeneracy of the universal nuclear genetic code.

[00110] In some embodiments, the isolated nucleic acid comprises a nucleotide sequence of SEQ ID NO: 20, and fragments thereof, that encode a polypeptide having a sequence of SEQ ID

NO: 7, or fragments thereof. The fragment of the nucleotide sequence of SEQ ID NO: 20 comprises at least one or more nucleotides from one or more of the following regions of SEQ ID NO: 20: nucleotides in SEQ ID NOs: SEQ ID NOs: 14, 15, 16, 17, 18, 19 or 20.

[00111] The invention further embraces other species, homologs of the Rel_{Mtb} DNA. Species homologs, sometimes referred to as "orthologs," in general, share at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% homology with Rel_{Mtb} DNA of the invention. Generally, percent sequence "homology" with respect to polynucleotides of the invention may be calculated as the percentage of nucleotide bases in the candidate sequence that are identical to nucleotides in the Rel_{Mtb} sequence set forth in sequences of SEQ ID NOs: SEQ ID NOs: 14, 15, 16, 17, 18, 19 or 20, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity.

[00112] The disclosure herein of full-length polynucleotides encoding Rel_{Mtb} polypeptides makes readily available to the worker of ordinary skill in the art other fragments of the full-length polynucleotide.

[00113] In some embodiments the present invention provides an isolated nucleic acid molecule comprising a sequence homologous to sequences of SEQ ID NOs: 14, 15, 16, 17, 18, 19 or 20, and fragments thereof. In some embodiments, the present invention provides an isolated nucleic acid molecule comprising a sequence of SEQ ID NOs: 14, 15, 16, 17, 18, 19 or 20, and fragments thereof.

[00114] As used in the present invention, fragments of Rel_{Mtb}-encoding polynucleotides comprise at least 10, and in some embodiments at least 12, 14, 16, 18, 20, 25, 50, 75, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, or 2000 consecutive nucleotides of a polynucleotide encoding Rel_{Mtb}. Preferably, fragment polynucleotides of the invention comprise sequences unique to the Rel_{Mtb}-encoding polynucleotide sequence, and therefore hybridize under highly stringent or moderately stringent conditions only (i.e., "specifically") to polynucleotides encoding Rel_{Mtb} (or fragments thereof). Sequences unique to polynucleotides of the invention are recognizable through sequence comparison to other known polynucleotides, and can be identified through use of alignment programs routinely utilized in the art, *e.g.*, those made available in public sequence databases. Such sequences also are recognizable from Southern hybridization analyses to determine the number of fragments of genomic DNA to which a polynucleotide will hybridize.

Polynucleotides of the invention can be labeled in a manner that permits their detection, including radioactive, fluorescent, and enzymatic labeling.

[00115] Fragment polynucleotides are particularly useful as probes for detection of full-length or fragments of Rel_{Mtb} polynucleotides. One or more polynucleotides can be included in kits that are used to detect the presence of a polynucleotide encoding Rel_{Mtb}, or used to detect variations in a polynucleotide sequence encoding Rel_{Mtb}.

[00116] The invention also embraces DNAs encoding Rel_{Mtb} polypeptides that hybridize under moderately stringent or high stringency conditions to the non-coding strand, or complement, of the polynucleotides set forth in sequences of SEQ ID NOs:14-20.

[00117] In some embodiments, highly stringent hybridization conditions are as follows: hybridization at 42□C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% Dextran sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1X SSC and 1% SDS. It is understood in the art that conditions of equivalent stringency can be achieved through variation of temperature and buffer, or salt concentration as described Ausubel et al. (Eds.), Protocols in Molecular Biology, John Wiley & Sons (1994), pp. 6.0.3 to 6.4.10. Modifications in hybridization conditions can be empirically determined or precisely calculated based on the length and the percentage of guanosine/cytosine (GC) base pairing of the probe. The hybridization conditions can be calculated as described in Sambrook, et al., (Eds.), Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York (1989), pp. 9.47 to 9.51.

[00118] The polynucleotide sequence information provided by the invention makes possible large-scale expression of the encoded polypeptide by techniques well known and routinely practiced in the art.

[00119] Vectors

[00120] Another aspect of the present invention is directed to vectors, or recombinant expression vectors, comprising any of the nucleic acid molecules described above. Vectors are used herein either to amplify DNA or RNA encoding Rel_{Mtb} and/or to express DNA which encodes Rel_{Mtb}. In some embodiments, vectors include, but are not limited to, plasmids, phages, cosmids, episomes, viral particles or viruses, and integratable DNA fragments (i.e., fragments integratable into the host genome by homologous recombination). In some embodiments viral particles include, but are not limited to, adenoviruses, baculoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses, and retroviruses. In some embodiments, expression vectors include, but are not limited to,

pcDNA3 (Invitrogen) and pSVL (Pharmacia Biotech). Other expression vectors include, but are not limited to, pSPORT™ vectors, pGEM™ vectors (Promega), pPROEXvectors™ (LTI, Bethesda, MD), Bluescript™ vectors (Stratagene), pQE™ vectors (Qiagen), pSE420™ (Invitrogen), and pYES2™(Invitrogen).

[00121] Expression constructs preferably comprise Rel_{Mtb}-encoding polynucleotides operatively linked to an endogenous or exogenous expression control DNA sequence and a transcription terminator. Expression control DNA sequences include promoters, enhancers, operators, and regulatory element binding sites generally, and are typically selected based on the expression systems in which the expression construct is to be utilized. In some embodiments, promoter and enhancer sequences are selected for the ability to increase gene expression, while operator sequences are selected for the ability to regulate gene expression. Expression constructs of the invention may also include sequences encoding one or more selectable markers that permit identification of host cells bearing the construct. Expression constructs may also include sequences that facilitate, and preferably promote, homologous recombination in a host cell. Constructs of the invention can also include sequences necessary for replication in a host cell.

[00122] Expression constructs are utilized for production of an encoded protein, but may also be utilized simply to amplify a Rel_{Mtb}-encoding polynucleotide sequence. In some embodiments, the vector is an expression vector wherein the polynucleotide of the invention is operatively linked to a polynucleotide comprising an expression control sequence. Autonomously replicating recombinant expression constructs such as plasmid and viral DNA vectors incorporating polynucleotides of the invention are also provided. In some embodiments expression vectors are replicable DNA constructs in which a DNA sequence encoding Rel_{Mtb} is operably linked or connected to suitable control sequences capable of effecting the expression of the Rel_{Mtb} in a suitable host. DNA regions are operably linked or connected when they are functionally related to each other. For example, a promoter is operably linked or connected to a coding sequence if it controls the transcription of the sequence. Amplification vectors do not require expression control domains, but rather need only the ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants. The need for control sequences in the expression vector will vary depending upon the host selected and the transformation method chosen. Generally, control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding and sequences which control the termination of transcription and translation.

[00123] In some embodiments, vectors contain a promoter that is recognized by the host organism. The promoter sequences of the present invention may be prokaryotic, eukaryotic or viral. Examples of suitable prokaryotic sequences include the PR and PL promoters of bacteriophage lambda (The bacteriophage Lambda, Hershey, A. D., Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1973), which is incorporated herein by reference in its entirety; Lambda II, Hendrix, R. W., Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1980), which is incorporated herein by reference in its entirety); the trp, recA, heat shock, and lacZ promoters of E. coli and the SV40 early promoter (Benoist et al. Nature, 1981, 290, 304-310, which is incorporated herein by reference in its entirety). Additional promoters include, but are not limited to, mouse mammary tumor virus, long terminal repeat of human immunodeficiency virus, maloney virus, cytomegalovirus immediate early promoter, Epstein Barr virus, Rous sarcoma virus, human actin, human myosin, human hemoglobin, human muscle creatine, and human metalothionein.

[00124] Additional regulatory sequences can also be included in vectors. In some embodiments, examples of suitable regulatory sequences are represented by the Shine-Dalgarno of the replicase gene of the phage MS-2 and of the gene cII of bacteriophage lambda. The Shine-Dalgarno sequence may be directly followed by DNA encoding Rel_{Mtb} or fragments thereof and result in the expression of the mature Rel_{Mtb} protein or fragments thereof.

[00125] Moreover, suitable expression vectors can include an appropriate marker that allows the screening of the transformed host cells. The transformation of the selected host is carried out using any one of the various techniques well known to the expert in the art and described in Sambrook et al., supra.

[00126] An origin of replication can also be provided either by construction of the vector to include an exogenous origin or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter may be sufficient. Alternatively, rather than using vectors which contain viral origins of replication, one skilled in the art can transform mammalian cells by the method of co-transformation with a selectable marker and Rel_{Mtb} DNA. An example of a suitable marker is dihydrofolate reductase (DHFR) or thymidine kinase (see, U.S. Patent No. 4,399,216).

[00127] Nucleotide sequences encoding Rel_{Mtb} may be recombined with vector DNA in accordance with conventional techniques, including blunt-ended or staggered-ended termini for ligation, restriction enzyme digestion to provide appropriate termini, filling in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and ligation with

appropriate ligases. Techniques for such manipulation are disclosed by Sambrook et al., supra and are well known in the art. Methods for construction of mammalian expression vectors are disclosed in, for example, Okayama et al., Mol. Cell. Biol., 1983, 3, 280, Cosman et al., Mol. Immunol., 1986, 23, 935, Cosman et al., Nature, 1984, 312, 768, EP-A-0367566, and WO 91/18982, each of which is incorporated herein by reference in its entirety.

Host Cells

[00128] According to some embodiments of the invention, host cells are provided, including prokaryotic and eukaryotic cells, comprising a polynucleotide of the invention (or vector of the invention) in a manner that permits expression of the encoded Rel_{Mtb} polypeptide or fragments thereof. Polynucleotides of the invention may be introduced into the host cell as part of a circular plasmid, or as linear DNA comprising an isolated protein coding region or a viral vector. Methods for introducing DNA into the host cell that are well known and routinely practiced in the art include transformation, transfection, electroporation, nuclear injection, or fusion with carriers such as liposomes, micelles, ghost cells, and protoplasts. Expression systems of the invention include bacterial, yeast, fungal, plant, insect, invertebrate, vertebrate, and mammalian cells systems. In some embodiments the bacterial system is M. smegmatis or *M. tuberculosis*.

[00129] The invention provides host cells that are transformed or transfected (stably or transiently) with polynucleotides of the invention or vectors of the invention. As stated above, such host cells are useful for amplifying the polynucleotides and also for expressing the Rel_{Mtb} polypeptide or fragments thereof encoded by the polynucleotide.

[00130] One of skill in the art can introduce nucleic acid molecules into mycobacteria (e.g. *M. tuberculosis* or *M. smegmatis*) using methods known to those skilled in the art (see, for example, *Proc Natl Acad Sci* U S A. 1991 Apr 15;88(8):3111-5; *Mol Microbiol.* 1989 Jan;3(1):29-34; and *Nutrition.* 1995 Sep-Oct;11(5 Suppl):670-3). These methods can be used *inter alia*, for expressing exogenous proteins and homologous recombination.

[00131] In some embodiments, the invention provides a method for producing a Rel_{Mtb} polypeptide (or fragment thereof) comprising the steps of growing a host cell of the invention in a nutrient medium and isolating the polypeptide or variant thereof from the cell or the medium.

[00132] According to some aspects of the present invention, transformed host cells having an expression vector comprising any of the nucleic acid molecules described above are provided. Expression of the nucleotide sequence occurs when the expression vector is introduced into an appropriate host cell. Suitable host cells for expression of the polypeptides of the invention include, but are not limited to, prokaryotes, yeast, and eukaryotes. If a prokaryotic expression

vector is employed, then the appropriate host cell would be any prokaryotic cell capable of expressing the cloned sequences. Suitable prokaryotic cells include, but are not limited to, bacteria of the genera Escherichia, Bacillus, Salmonella, Pseudomonas, Streptomyces, and Staphylococcus. In some embodiments the prokaryotic cell is M. smegmatis or *M. tuberculosis*.

[00133] If an eukaryotic expression vector is employed, then the appropriate host cell would be any eukaryotic cell capable of expressing the cloned sequence. In some embodiments, eukaryotic cells are cells of higher eukaryotes. Suitable eukaryotic cells include, but are not limited to, non-human mammalian tissue culture cells and human tissue culture cells. In some embodiments, host cells include, but are not limited to, insect cells, HeLa cells, Chinese hamster ovary cells (CHO cells), African green monkey kidney cells (COS cells), human HEK-293 cells, and murine 3T3 fibroblasts. Propagation of such cells in cell culture has become a routine procedure (see, Tissue Culture, Academic Press, Kruse and Patterson, eds. (1973), which is incorporated herein by reference in its entirety).

[00134] In addition, a yeast host may be employed as a host cell. Preferred yeast cells include, but are not limited to, the genera Saccharomyces, Pichia, and Kluveromyces. Preferred yeast hosts are S. cerevisiae and P. pastoris. Preferred yeast vectors can contain an origin of replication sequence from a 2T yeast plasmid, an autonomously replication sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination, and a selectable marker gene. Shuttle vectors for replication in both yeast and E. coli are also included herein.

[00135] Alternatively, insect cells may be used as host cells. In a preferred embodiment, the polypeptides of the invention are expressed using a baculovirus expression system (see, Luckow et al., Bio/Technology, 1988, 6, 47, Baculovirus Expression Vectors: A Laboratory Manual, O'Rielly et al. (Eds.), W.H. Freeman and Company, New York, 1992, and U.S. Patent No. 4,879,236, each of which is incorporated herein by reference in its entirety). In addition, the MAXBAC™ complete baculovirus expression system (Invitrogen) can, for example, be used for production in insect cells.

[00136] Host cells of the invention are a valuable source of immunogen for development of antibodies specifically immunoreactive with Rel_{Mtb}. Host cells of the invention are also useful in methods for the large-scale production of Rel_{Mtb} polypeptides wherein the cells are grown in a suitable culture medium and the desired polypeptide products are isolated from the cells, or from the medium in which the cells are grown, by purification methods known in the art, *e.g.*, conventional chromatographic methods including immunoaffinity chromatography, protein

affinity chromatography, hydrophobic interaction chromatography, lectin affinity chromatography, size exclusion filtration, cation or anion exchange chromatography, high pressure liquid chromatography (HPLC), reverse phase HPLC, and the like. Still other methods of purification include those methods wherein the desired protein is expressed and purified as a fusion protein having a specific tag, label, or chelating moiety that is recognized by a specific binding partner or agent. The purified protein can be cleaved to yield the desired protein, or can be left as an intact fusion protein. Cleavage of the fusion component may produce a form of the desired protein having additional amino acid residues as a result of the cleavage process.

[00137] Knowledge of Rel_{Mtb} DNA sequences allows for modification of cells to permit, or increase, expression of endogenous Rel_{Mtb}. Cells can be modified (*e.g.*, by homologous recombination) to provide increased expression by replacing, in whole or in part, the naturally occurring Rel_{Mtb} promoter with all or part of a heterologous promoter so that the cells express Rel_{Mtb} at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to endogenous Rel_{Mtb} encoding sequences. (See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955.) It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, ada, dhfr, and the multifunctional CAD gene which encodes carbamoyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the Rel_{Mtb} coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the Rel_{Mtb} coding sequences in the cells.

Knock-outs

[00138] The DNA sequence information provided by the present invention also makes possible the development of bacteria that fail to express functional Rel_{Mtb} or that express a variant of Rel_{Mtb}. Such models are useful for studying the *in vivo* activities of Rel_{Mtb} and modulators of Rel_{Mtb} activity. Knock-outs can also be generated by using a mutagen to generate mutants of bacteria that don't express a specific gene or genes.

Antisense

[00139] Also made available by the invention are anti-sense polynucleotides that recognize and hybridize to polynucleotides encoding Rel_{Mtb}. Full-length and fragment anti-sense polynucleotides are provided. Fragment antisense molecules of the invention include (i) those that specifically recognize and hybridize to Rel_{Mtb} RNA (as determined by sequence comparison

of DNA encoding Rel_{Mtb} to DNA encoding other known molecules). Identification of sequences unique to Rel_{Mtb} encoding polynucleotides can be deduced through use of any publicly available sequence database, and/or through use of commercially available sequence comparison programs. After identification of the desired sequences, isolation through restriction digestion or amplification using any of the various polymerase chain reaction techniques well known in the art can be performed. Anti-sense polynucleotides are particularly relevant to regulating expression of Rel_{Mtb} by those cells expressing Rel_{Mtb} mRNA.

[00140] Antisense nucleic acids (*e.g.*10 to 30 base-pair oligonucleotides) capable of specifically binding to Rel_{Mtb} expression control sequences or Rel_{Mtb} RNA are introduced into cells (*e.g.*, by a viral vector or colloidal dispersion system such as a liposome). The antisense nucleic acid binds to the Rel_{Mtb} target nucleotide sequence in the cell and prevents transcription and/or translation of the target sequence. Phosphorothioate and methylphosphonate antisense oligonucleotides are contemplated for therapeutic use by the invention. The antisense oligonucleotides may be further modified by adding poly-L-lysine, transferrin polylysine, or cholesterol moieties at their 5' end. Suppression of Rel_{Mtb} expression at either the transcriptional or translational level is useful to generate cellular or animal models for diseases/conditions characterized by Rel_{Mtb} expression.

[00141] Antisense oligonucleotides, or fragments of sequences of SEQ ID NOS: SEQ ID NOs: 14, 15, 16, 17, 18, 19 or 20, or sequences complementary or homologous thereto, derived from the nucleotide sequences of the present invention encoding Rel_{Mtb} are useful as diagnostic tools for probing gene expression in various tissues. For example, tissue can be probed in situ with oligonucleotide probes carrying detectable groups by conventional autoradiography techniques to investigate native expression of this protein or pathological conditions relating thereto. Antisense oligonucleotides can be directed to domains or regulatory regions of sequences of SEQ ID NOS: SEQ ID NOs: 14, 15, 16, 17, 18, 19 or 20, or mRNA corresponding thereto, including, but not limited to, the initiation codon, TATA box, enhancer sequences, and the like.

Polypeptides

[00142] The invention provides purified and isolated Rel_{Mtb} polypeptides encoded by a polynucleotide of the invention. In some embodiments the Rel_{Mtb} polypeptide comprises the amino acid sequence set out in sequences of SEQ ID NOs: 7-13, or fragments thereof comprising an epitope specific to the polypeptide. By "epitope specific to" is meant a portion of the Rel_{Mtb}

protein that is recognizable by an antibody that is specific for the Rel_{Mtb}, as defined in detail below.

[00143] Although the sequences provided are particular *M. tuberculosis* sequences, the invention is intended to include within its scope other bacterial forms of Rel_{Mtb}.

[00144] The invention also embraces polypeptides that have at least 99%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55% or at least 50% identity and/or homology to the preferred polypeptides of the invention. Percent amino acid sequence "identity" with respect to the preferred polypeptide of the invention is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the residues in the Rel_{Mtb} sequence after aligning both sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Percent sequence "homology" with respect to the preferred polypeptide of the invention is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the residues in the Rel_{Mtb} sequence after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and also considering any conservative substitutions as part of the sequence identity. Homologous peptides can also refer to peptides that have a similar function to Rel_{Mtb} as well as sequence homology.

[00145] In some aspects, percent homology is calculated as the percentage of amino acid residues in the smaller of two sequences which align with identical amino acid residue in the sequence being compared, when four gaps in a length of 100 amino acids may be introduced to maximize alignment (Dayhoff, in Atlas of Protein Sequence and Structure, Vol. 5, p. 124, National Biochemical Research Foundation, Washington, D.C. (1972), incorporated herein by reference).

[00146] Polypeptides of the invention may be isolated from natural cell sources or may be chemically synthesized, but can be produced by recombinant procedures involving host cells of the invention. Use of mammalian host cells can be used to provide for such post-translational modifications (*e.g.*, glycosylation, truncation, lipidation, and phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. Glycosylated and non-glycosylated forms of Rel_{Mtb} polypeptides are embraced by the invention.

[00147] The invention also embraces variant (or analog) Rel_{Mtb} polypeptides. In one example, insertion variants are provided wherein one or more amino acid residues supplement a Rel_{Mtb} amino acid sequence. Insertions may be located at either or both termini of the protein, or

may be positioned within internal regions of the Rel_{Mtb} amino acid sequence. Insertional variants with additional residues at either or both termini can include, for example, fusion proteins and proteins including amino acid tags or labels.

[00148] Insertion variants include Rel_{Mtb} polypeptides wherein one or more amino acid residues are added to a Rel_{Mtb} acid sequence or to a biologically active fragment thereof.

[00149] Variant products of the invention also include mature Rel_{Mtb} products, i.e., Rel_{Mtb} products wherein leader or signal sequences are removed, with additional amino terminal residues. The additional amino terminal residues may be derived from another protein, or may include one or more residues that are not identifiable as being derived from specific proteins. Rel_{Mtb} products with an additional methionine residue at position -1 (Met-1-Rel_{Mtb}) are contemplated, as are variants with additional methionine and lysine residues at positions -2 and -1 (Met-2-Lys-1- Rel_{Mtb}). Variants of Rel_{Mtb} with additional Met, Met-Lys, Lys residues (or one or more basic residues in general) are particularly useful for enhanced recombinant protein production in bacterial host cells.

[00150] The invention also embraces Rel_{Mtb} variants having additional amino acid residues that result from use of specific expression systems. For example, use of commercially available vectors that express a desired polypeptide as part of a glutathione-S-transferase (GST) fusion product provides the desired polypeptide having an additional glycine residue at position -1 after cleavage of the GST component from the desired polypeptide. Variants that result from expression in other vector systems are also contemplated.

[00151] Insertional variants also include fusion proteins wherein the amino terminus and/or the carboxy terminus of Rel_{Mtb} is/are fused to another polypeptide.

[00152] In some aspects, the invention provides deletion variants wherein one or more amino acid residues in a Rel_{Mtb} polypeptide are removed. Deletions can be effected at one or both termini of the Rel_{Mtb} polypeptide, or with removal of one or more non-terminal amino acid residues of Rel_{Mtb}. Deletion variants, therefore, include all fragments of a Rel_{Mtb} polypeptide.

[00153] The invention also embraces polypeptide fragments of sequences of SEQ ID NO:7, wherein the fragments maintain biological (*e.g.*, hydrolysis and/or synthesis activity) and/or immunological properties of a Rel_{Mtb} polypeptide.

[00154] In some embodiments of the invention, an isolated nucleic acid molecule comprises a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence homologous to sequences of SEQ ID NOs:7-13, and fragments thereof, wherein the nucleic acid molecule encoding at least a portion of Rel_{Mtb}.

[00155] As used in the present invention, polypeptide fragments comprise at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 200, 225, 250, 300, 350, 400, 450, or 500 consecutive amino acids of sequences of SEQ ID NO:7. In some embodiments, polypeptide fragments display antigenic properties unique to, or specific for, Rel_{Mtb}. Fragments of the invention having the desired biological and immunological properties can be prepared by any of the methods well known and routinely practiced in the art. As discussed above, in some embodiments, the fragments have functional activity.

[00156] In still another aspect, the invention provides substitution variants of Rel_{Mtb} polypeptides. Substitution variants include those polypeptides wherein one or more amino acid residues of a Rel_{Mtb} polypeptide are removed and replaced with alternative residues. In one aspect, the substitutions are conservative in nature; however, the invention embraces substitutions that are also non-conservative. Conservative substitutions for this purpose may be defined as set out in Tables 1, 2, or 3 below.

[00157] Variant polypeptides include those wherein conservative substitutions have been introduced by modification of polynucleotides encoding polypeptides of the invention. Amino acids can be classified according to physical properties and contribution to secondary and tertiary protein structure. A conservative substitution is recognized in the art as a substitution of one amino acid for another amino acid that has similar properties. Exemplary conservative substitutions are set out in Table 1 (from WO 97/09433, page 10, published March 13, 1997 (PCT/GB96/02197, filed 9/6/96), immediately below.

Table 1

Conservative Substitutions I

SIDE CHAIN CHARACTERISTIC	AMINO ACID
Aliphatic	
Non-polar	G A P I L V
Polar - uncharged	C S T M N Q
Polar - charged	D E K R
Aromatic	H F W Y
Other	N Q D E

[00158] Alternatively, conservative amino acids can be grouped as described in Lehninger, [Biochemistry, Second Edition; Worth Publishers, Inc. NY, NY (1975), pp.71-77] as set out in Table 2, below.

Table 2
Conservative Substitutions II

SIDE CHAIN CHARACTERISTIC	AMINO ACID
Non-polar (hydrophobic)	
A. Aliphatic:	A L I V P
B. Aromatic:	F W
C. Sulfur-containing:	M
D. Borderline:	G
Uncharged-polar	
A. Hydroxyl:	S T Y
B. Amides:	N Q
C. Sulfhydryl:	C
D. Borderline:	G
Positively Charged (Basic):	K R H
Negatively Charged (Acidic):	D E

[00159] As still another alternative, exemplary conservative substitutions are set out in Table 3, below.

Table 3
Conservative Substitutions III

Original Residue	Exemplary Substitution
Ala (A)	Val, Leu, Ile
Arg (R)	Lys, Gln, Asn
Asn (N)	Gln, His, Lys, Arg
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
His (H)	Asn, Gln, Lys, Arg
Ile (I)	Leu, Val, Met, Ala, Phe,
Leu (L)	Ile, Val, Met, Ala, Phe
Lys (K)	Arg, Gln, Asn
Met (M)	Leu, Phe, Ile
Phe (F)	Leu, Val, Ile, Ala
Pro (P)	Gly
Ser (S)	Thr
Thr (T)	Ser

Trp (W)	Tyr
Tyr (Y)	Trp, Phe, Thr, Ser
Val (V)	Ile, Leu, Met, Phe, Ala

[00160] It should be understood that the definition of polypeptides of the invention is intended to include polypeptides bearing modifications other than insertion, deletion, or substitution of amino acid residues. By way of example, the modifications may be covalent in nature, and include for example, chemical bonding with polymers, lipids, other organic, and inorganic moieties. Such derivatives may be prepared to increase circulating half-life of a polypeptide, or may be designed to improve the targeting capacity of the polypeptide for desired cells, tissues, or organs. Similarly, the invention further embraces Rel_{Mtb} polypeptides that have been covalently modified to include one or more water-soluble polymer attachments such as polyethylene glycol, polyoxyethylene glycol, or polypropylene glycol. Variants that display activity properties of native Rel_{Mtb} and are expressed at higher levels, as well as variants that provide for constitutively active Rel_{Mtb}, can be useful in assays of the invention; the variants are also useful in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant Rel_{Mtb} activity.

[00161] In a related embodiment, the present invention provides compositions comprising purified polypeptides of the invention. In some embodiments, compositions comprise, in addition to the polypeptide of the invention, a pharmaceutically acceptable (i.e., sterile and non-toxic) liquid, semisolid, or solid diluent that serves as a pharmaceutical vehicle, excipient, or medium. Any diluent known in the art may be used. Exemplary diluents include, but are not limited to, water, saline solutions, polyoxyethylene sorbitan monolaurate, magnesium stearate, methyl- and propylhydroxybenzoate, talc, alginates, starches, lactose, sucrose, dextrose, sorbitol, mannitol, glycerol, calcium phosphate, mineral oil, and cocoa butter.

[00162] Variants that display activity properties of native Rel_{Mtb} and are expressed at higher levels, as well as variants that provide for constitutively active Rel_{Mtb}, are particularly useful in assays of the invention; the variants are also useful in assays of the invention and in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant Rel_{Mtb} activity.

[00163] Also comprehended by the present invention are antibodies (*e.g.*, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR sequences which

specifically recognize a polypeptide of the invention) specific for Rel_{Mtb} or fragments thereof. In some embodiments antibodies of the invention are human antibodies that are produced and identified according to methods described in WO93/11236, published June 20, 1993, which is incorporated herein by reference in its entirety. Antibody fragments, including Fab, Fab', F(ab')2, and Fv, are also provided by the invention. The term "specific for," when used to describe antibodies of the invention, indicates that the variable regions of the antibodies of the invention recognize and bind Rel_{Mtb} polypeptides exclusively. It will be understood that specific antibodies may also interact with other proteins (for example, S. aureus protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and, in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow et al. (Eds.), Antibodies A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the Rel_{Mtb} polypeptides of the invention are also contemplated, provided that the antibodies are specific for Rel_{Mtb} polypeptides. Antibodies of the invention can be produced using any method well known and routinely practiced in the art.

[00164] The invention provides an antibody that is specific for the Rel_{Mtb} of the invention. Antibody specificity is described in greater detail below. However, it should be emphasized that antibodies that can be generated from polypeptides that have previously been described in the literature and that are capable of fortuitously cross-reacting with Rel_{Mtb} (*e.g.*, due to the fortuitous existence of a similar epitope in both polypeptides) are considered "cross-reactive" antibodies. Such cross-reactive antibodies are not antibodies that are "specific" for Rel_{Mtb}. The determination of whether an antibody is specific for Rel_{Mtb} or is cross-reactive with another known protein is made using any of several assays, such as Western blotting assays, that are well known in the art. In some embodiments, the antibodies are specific for the polypeptides encoded by the nucleic acid molecules wherein the nucleic acid molecule comprises SEQ ID NOs. 14, 15, 16, 17, 18, 19, or 20. In some embodiments, the antibodies are specific for the polypeptides having a sequence of SEQ ID NOs. 7, 8, 9, 10, 11, 12, or 13.

[00165] In some embodiments, the invention provides monoclonal antibodies. Hybridomas that produce such antibodies also are intended as aspects of the invention. In yet another variation, the invention provides a humanized antibody. Humanized antibodies are useful for *in vivo* therapeutic indications.

[00166] In another variation, the invention provides a cell-free composition comprising polyclonal antibodies, wherein at least one of the antibodies is an antibody of the invention specific for Rel_{Mtb}. Antisera isolated from an animal is an exemplary composition, as is a composition comprising an antibody fraction of an antisera that has been resuspended in water or in another diluent, excipient, or carrier.

[00167] In still another related embodiment, the invention provides an anti-idiotypic antibody specific for an antibody that is specific for Rel_{Mtb}.

[00168] It is well known that antibodies contain relatively small antigen binding domains that can be isolated chemically or by recombinant techniques. Such domains are useful Rel_{Mtb} binding molecules themselves, and also may be reintroduced into human antibodies, or fused to toxins or other polypeptides. Thus, in still another embodiment, the invention provides a polypeptide comprising a fragment of a Rel_{Mtb}-specific antibody, wherein the fragment and the polypeptide bind to the Rel_{Mtb}. By way of non-limiting example, the invention provides polypeptides that are single chain antibodies and CDR-grafted antibodies.

[00169] Non-human antibodies may be humanized by any of the methods known in the art. In one method, the non-human CDRs are inserted into a human antibody or consensus antibody framework sequence. Further changes can then be introduced into the antibody framework to modulate affinity or immunogenicity.

[00170] Antibodies of the invention are useful for, *e.g.*, therapeutic purposes (by modulating activity of Rel_{Mtb}), diagnostic purposes to detect or quantitate Rel_{Mtb}, and purification of Rel_{Mtb}. Kits comprising an antibody of the invention for any of the purposes described herein are also comprehended. In general, a kit of the invention also includes a control antigen for which the antibody is immunospecific.

Compositions

[00171] Expression of Rel_{Mtb} is likely connected to the pathogenesis of tuberculosis. It is contemplated that, preventing the expression of, or inhibiting the activity of, Rel_{Mtb} will be useful in treating tuberculosis. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of Rel_{Mtb}.

[00172] Other aspects of the present invention are directed to compositions, including pharmaceutical compositions, comprising any of the nucleic acid molecules or recombinant expression vectors described above and an acceptable carrier or diluent. In some embodiments, the carrier or diluent is pharmaceutically acceptable. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, A. Osol, a standard reference text in this

field, which is incorporated herein by reference in its entirety. In some embodiments, examples of such carriers or diluents include, but are not limited to, water, saline, Ringer's solution, dextrose solution, and 5% human serum albumin. Liposomes and nonaqueous vehicles such as fixed oils may also be used. The formulations are sterilized by commonly used techniques.

[00173] Also within the scope of the invention are compositions comprising polypeptides, polynucleotides, or antibodies of the invention that have been formulated with, *e.g.*, a pharmaceutically acceptable carrier.

[00174] The invention also provides methods of using antibodies of the invention. For example, the invention provides a method for modulating the activity of a Rel_{Mtb} comprising the step of contacting the Rel_{Mtb} with an antibody specific for the Rel_{Mtb}, under conditions wherein the antibody binds the protein.

Formulation, Dose and Administration

[00175] Suitable formulations for administration of a composition of the invention to a subject include aqueous and non-aqueous sterile injection solutions which can contain anti-oxidants, buffers, bacteriostats, antibacterial and antifungal agents that render the formulation isotonic with the bodily fluids of the intended patient (*e.g.*, sugars, salts, and polyalcohols), suspending agents and thickening agents. The formulations can be presented in unit-dose or multi-dose containers, for example sealed ampoules and vials, and can be stored in a frozen or freeze-dried (lyophilized) condition requiring only the addition of sterile liquid carrier immediately prior to use.

[00176] Compositions useful for injection into a patient include sterile aqueous solutions or dispersions, and sterile powder for the preparation of sterile injectable solutions or dispersions. An injectable composition should be fluid to the extent that administration via a syringe is readily performed. Suitable solvents include water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol), and mixtures thereof. Fluidity can be maintained, for example, by the use of a coating such as lecithin and/or by minimization of particle size.

[00177] An inhibitor of the present invention can be administered to a patient intratumorally, peritumorally, systemically, parenterally (*e.g.*, intravenous injection, intra-muscular injection, intra-arterial injection, and infusion techniques), orally, transdermally (topically), intranasally (inhalation), and intramucosally. A delivery method is selected based on considerations such as the type of the type of carrier or vector, therapeutic efficacy of the composition, location of target area, and the condition to be treated.

[00178] In some embodiments, the inhibitor is administered orally or intravenously to the patient. In some embodiments, the inhibitor is administered systemically. In some embodiments the inhibitor is administered locally to affected regions.

[00179] The present invention provides that an effective amount of a non-pathogenic virus is administered to a subject. The term "effective amount" is used herein to describe an amount of an inhibitor sufficient to elicit anti-tuberculosis activity. In some embodiments the anti-tuberculosis activity is through inhibition of the electron transport system of *M. tuberculosis*. In some embodiments, the anti-tuberculosis activity is through inhibition of the Rel protein activity of *M. tuberculosis*.

[00180] In some embodiments, an "effective amount" refers to the amount of a therapeutic that is effective in an *in vitro* assay in inhibiting *M. tuberculosis* cell growth, inhibiting *M. tuberculosis* cell respiration, inhibiting *M. tuberculosis* cell replication, and the like. In some embodiments, an "effective amount" inhibits *M. tuberculosis* cell growth, *M. tuberculosis* cell respiration, *M. tuberculosis* cell replication, or combinations thereof at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 2-fold, at least 5-fold, at least 10-fold, or at least 100-fold.

[00181] Actual dosage levels of active ingredients in a therapeutic composition of the invention can be varied so as to administer an amount of the composition that is effective to achieve the desired therapeutic response for a particular subject. Administration regimens can also be varied as required to elicit the desired activity. A single injection or multiple injections can be used. The selected dosage level and regimen will depend upon a variety of factors including the activity of the therapeutic composition, formulation, the route of administration, combination with other drugs or treatments, the disease or disorder to be treated, and the physical condition and prior medical history of the subject being treated. Determination and adjustment of an effective amount or dose, as well as evaluation of when and how to make such adjustments, are known to those of ordinary skill in the art of medicine.

[00182] As used herein, the term "pharmaceutical solid dosage forms" refers to a final solid pharmaceutical product. The term "pharmaceutical solid dosage form" includes, but is not limited to, tablets, caplets, beads, and capsules (including both hard shell capsules and soft gelatin capsules).

[00183] As used herein, the term "pharmaceutical liquid dosage form" includes liquid forms in which the compounds and compositions of the present invention can be incorporated for administration orally or by injection, and include aqueous solutions, suitably flavored syrups,

aqueous or oil suspensions, and flavored emulsions with edible oils such as cottonseed oil, sesame oil, coconut oil, or peanut oil, as well as elixirs and similar pharmaceutical vehicles.

[00184] For additional guidance regarding formulation, dose and administration regimen, see Berkow et al. (1997) The Merck Manual of Medical Information, Home ed. Merck Research Laboratories, Whitehouse Station, New Jersey; Goodman et al. (1996) Goodman & Gilman's the Pharmacological Basis of Therapeutics, 9th ed. McGraw-Hill Health Professions Division, New York; Ebadi (1998) CRC Desk Reference of Clinical Pharmacology. CRC Press, Boca Raton, Florida; Katzung (2001) Basic & Clinical Pharmacology, 8th ed. Lange Medical Books/McGraw-Hill Medical Pub. Division, New York; Remington et al. (1975) Remington's Pharmaceutical Sciences, 15th ed. Mack Pub. Co., Easton, Pennsylvania; Speight et al. (1997) Avery's Drug Treatment: A Guide to the Properties, Choice, Therapeutic Use and Economic Value of Drugs in Disease Management, 4th ed. Adis International, Auckland / Philadelphia, Pennsylvania.

[00185] In some embodiments, compositions are tested *in vitro* or *in vivo* assays in order to determine an "effective amount." For example, in methods disclosed herein for causing cell death, assays suitable include, without limitation, *in vitro* cell viability assays, including the TUNEL assay or other fluorescent based assays such as Cell-Titer Blue (Promega Corp).

[00186] The inhibitors of the present invention may further comprise one or more adjuvants which include, but are not limited to: (1) aluminum salts (alum), (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides or bacterial cell wall components), such as for example (a) MF59 containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP, and (c) RibiTM. adjuvant system (RAS), (Ribi Immunochem, Hamilton, Mont.) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS); (3) saponin adjuvants, (4) Complete Freunds Adjuvant (CFA) and Incomplete Freunds Adjuvant (IFA); (5) cytokines, such as interleukins (IL-1, IL-2, etc.), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc.; (6) detoxified mutants of a bacterial ADP-ribosylating toxin, and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition. Muramyl peptides include, without limitation, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acteyl-normuramyl-L-alanyl-D-isogluatme (nor-MDP), N-acetylmuramyl-L-

alanyl-D-isogluatminyl-L-alanine-2-(1'-2'-dipahitoyl-sn- glycero-3-huydroxyphosphoryloxy)-ethylamine (MTP-PE), and the like.

Kits

[00187] The present invention is also directed to kits, including pharmaceutical kits. The kits can comprise any of the nucleic acid molecules described above, any of the polypeptides described above, or any antibody which binds to a polypeptide of the invention as described above, as well as a negative control. The kit preferably comprises additional components, such as, for example, instructions, solid support, reagents helpful for quantification, and the like.

[00188] In some aspects, the invention features methods for detection of a polypeptide in a sample as a diagnostic tool for diseases or disorders, wherein the method comprises the steps of: (a) contacting the sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of a polypeptide having sequences of SEQ ID NO:7, 8, 9, 10, 11, 12, or 13, said probe comprising the nucleic acid sequence encoding the polypeptide, fragments thereof, and the complements of the sequences and fragments; and (b) detecting the presence or amount of the probe:target region hybrid as an indication of the disease. In some embodiments the disease is tuberculosis.

[00189] Kits may be designed to detect either expression of polynucleotides encoding Rel_{Mtb} expressed in the lungs or the Rel_{Mtb} proteins themselves in order to identify tissue as being from infected tissue. For example, oligonucleotide hybridization kits can be provided which include a container having an oligonucleotide probe specific for the Rel_{Mtb}-specific DNA and optionally, containers with positive and negative controls and/or instructions. Similarly, PCR kits can be provided which include a container having primers specific for the Rel_{Mtb}-specific sequences, DNA and optionally, containers with size markers, positive and negative controls and/or instructions.

[00190] Hybridization conditions should be such that hybridization occurs only with the genes in the presence of other nucleic acid molecules. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having 1 or 2 mismatches out of 20 contiguous nucleotides. Such conditions are defined supra.

[00191] Alternatively, immunoassay kits can be provided which have containers container having antibodies specific for the Rel_{Mtb}-protein and optionally, containers with positive and negative controls and/or instructions.

[00192] Kits may also be provided useful in the identification of Rel_{Mtb} binding partners such as natural ligands or modulators (agonists or antagonists). Substances useful for treatment of disorders or diseases preferably show positive results in one or more *in vitro* assays for an activity corresponding to treatment of the disease or disorder in question. Substances that modulate the activity of the polypeptides preferably include, but are not limited to, antisense oligonucleotides, agonists and antagonists, and inhibitors of protein kinases.

Methods of inducing immune response

[00193] Another aspect of the present invention is directed to methods of inducing an immune response in a mammal against a polypeptide of the invention by administering to the mammal an amount of the polypeptide sufficient to induce an immune response. The amount will be dependent on the animal species, size of the animal, and the like but can be determined by those skilled in the art.

Methods of identifying compounds that bind Rel_{Mtb} or fragments thereof

[00194] The invention also provides assays to identify compounds that bind Rel_{Mtb}. One such assay comprises the steps of: (a) contacting a composition comprising a Rel_{Mtb} with a compound suspected of binding Rel_{Mtb}; and (b) measuring binding between the compound and Rel_{Mtb}. In some embodiments, the Rel_{Mtb} comprises only the domains that are necessary for synthesis activity or hydrolysis activity. In some embodiments, the composition comprises a cell expressing Rel_{Mtb}. The binding may be measured directly, *e.g.*, by using a labeled compound, or may be measured indirectly by several techniques, including measuring intracellular activity of Rel_{Mtb} induced by the compound (or measuring changes in the level of Rel_{Mtb} activity). Following steps (a) and (b), compounds identified as binding Rel_{Mtb} may be tested in other assays including, but not limited to, *in vivo* models, to confirm or quantitate binding to Rel_{Mtb}.

[00195] Specific binding molecules, including natural ligands and synthetic compounds, can be identified or developed using isolated or recombinant Rel_{Mtb} products, Rel_{Mtb} fragments as described herein, or in some embodiments, cells expressing such products or fragments. Binding partners can be useful for purifying Rel_{Mtb} products and detection or quantification of Rel_{Mtb} products in fluid and tissue samples using known immunological procedures. Binding molecules are also manifestly useful in modulating (i.e., blocking, inhibiting or stimulating) biological activities of Rel_{Mtb}, especially those activities involved in hydrolysis, synthesis, or transcription.

[00196] The DNA and amino acid sequence information provided by the present invention also makes possible identification of binding partner compounds with which a Rel_{Mtb} polypeptide (or fragment thereof) or polynucleotide will interact. Methods to identify binding

partner compounds include solution assays, *in vitro* assays wherein Rel_{Mtb} polypeptides are immobilized, and cell-based assays. Identification of binding partner compounds of Rel_{Mtb} polypeptides provides candidates for therapeutic or prophylactic intervention in pathologies associated with Rel_{Mtb} normal and aberrant biological activity.

[00197] The invention includes several assay systems for identifying Rel_{Mtb} binding partners. In solution assays, methods of the invention comprise the steps of (a) contacting a Rel_{Mtb} polypeptide with one or more candidate binding partner compounds and (b) identifying the compounds that bind to the Rel_{Mtb} polypeptide. Identification of the compounds that bind the Rel_{Mtb} polypeptide can be achieved by isolating the Rel_{Mtb} polypeptide/binding partner complex, and separating the binding partner compound from the Rel_{Mtb} polypeptide. An additional step of characterizing the physical, biological, and/or biochemical properties of the binding partner compound is also comprehended in another embodiment of the invention, wherein compounds identified as binding Rel_{Mtb} may be tested in other assays including, but not limited to, *in vivo* models, to confirm or quantitate binding to Rel_{Mtb}. In one aspect, the Rel_{Mtb} polypeptide/binding partner complex is isolated using an antibody immunospecific for either the Rel_{Mtb} polypeptide or the candidate binding partner compound.

[00198] In still other embodiments, either the Rel_{Mtb} polypeptide or the candidate binding partner compound comprises a label or tag that facilitates its isolation, and methods of the invention to identify binding partner compounds include a step of isolating the Rel_{Mtb} polypeptide/binding partner complex through interaction with the label or tag. An exemplary tag of this type is a poly-histidine sequence, generally around six histidine residues, that permits isolation of a compound so labeled using nickel chelation. Other labels and tags, such as the FLAG® tag (Eastman Kodak, Rochester, NY), well known and routinely used in the art, are embraced by the invention.

[00199] In one variation of an *in vitro* assay, the invention provides a method comprising the steps of (a) contacting an immobilized Rel_{Mtb} polypeptide with a candidate binding partner compound and (b) detecting binding of the candidate compound to the Rel_{Mtb} polypeptide. In an alternative embodiment, the candidate binding partner compound is immobilized and binding of Rel_{Mtb} is detected. Immobilization is accomplished using any of the methods well known in the art, including covalent bonding to a support, a bead, or a chromatographic resin, as well as non-covalent, high affinity interactions such as antibody binding, or use of streptavidin/biotin binding wherein the immobilized compound includes a biotin moiety. Detection of binding can be accomplished (i) using a radioactive label on the compound that is not immobilized, (ii) using of

a fluorescent label on the non-immobilized compound, (iii) using an antibody immunospecific for the non-immobilized compound, (iv) using a label on the non-immobilized compound that excites a fluorescent support to which the immobilized compound is attached, as well as other techniques well known and routinely practiced in the art.

[00200] The invention also provides cell-based assays to identify binding partner compounds of a Rel_{Mtb} polypeptide. In one embodiment, the invention provides a method comprising the steps of contacting a Rel_{Mtb} polypeptide expressed on the surface of a cell with a candidate binding partner compound and detecting binding of the candidate binding partner compound to the Rel_{Mtb} polypeptide. In a preferred embodiment, the detection comprises detecting a calcium flux or other physiological event in the cell caused by the binding of the molecule.

[00201] Another aspect of the present invention is directed to methods of identifying compounds that bind to either Rel_{Mtb} or nucleic acid molecules encoding Rel_{Mtb}, comprising contacting Rel_{Mtb}, or a nucleic acid molecule encoding the same, with a compound, and determining whether the compound binds Rel_{Mtb} or a nucleic acid molecule encoding the same. Binding can be determined by binding assays which are well known to the skilled artisan, including, but not limited to, gel-shift assays, Western blots, radiolabeled competition assay, phage-based expression cloning, co-fractionation by chromatography, co-precipitation, cross linking, interaction trap/two-hybrid analysis, southwestern analysis, ELISA, and the like, which are described in, for example, Current Protocols in Molecular Biology, 1999, John Wiley & Sons, NY, which is incorporated herein by reference in its entirety. The compounds to be screened include (which may include compounds which are suspected to bind Rel_{Mtb}, or a nucleic acid molecule encoding the same), but are not limited to, extracellular, intracellular, biologic or chemical origin. The methods of the invention also embrace ligands, a fluorescence label, a chemiluminescent label, an enzymatic label and an immunogenic label. Modulators falling within the scope of the invention include, but are not limited to, non-peptide molecules such as non-peptide mimetics, non-peptide allosteric effectors, and peptides. The Rel_{Mtb} polypeptide or polynucleotide employed in such a test may either be free in solution, attached to a solid support, borne on a cell surface or located intracellularly or associated with a portion of a cell. One skilled in the art can, for example, measure the formation of complexes between Rel_{Mtb} and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between Rel_{Mtb} and its substrate caused by the compound being tested.

[00202] In another embodiment of the invention, high throughput screening for compounds having suitable binding affinity to Rel_{Mtb} is employed. Briefly, large numbers of different test

compounds are synthesized on a solid substrate. The peptide test compounds are contacted with Rel_{Mtb} and washed. Bound Rel_{Mtb} is then detected by methods well known in the art. Purified polypeptides of the invention can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the protein and immobilize it on the solid support.

[00203] Generally, an expressed Rel_{Mtb} can be used for HTS binding assays in conjunction with a defined partner. The identified peptide is labeled with a suitable radioisotope, including, but not limited to, 125I, 3H, 35S or 32P, by methods that are well known to those skilled in the art. Alternatively, the peptides may be labeled by well-known methods with a suitable fluorescent derivative (Baindur et al., Drug Dev. Res., 1994, 33, 373-398; Rogers, Drug Discovery Today, 1997, 2, 156-160). Radioactive ligand specifically bound to the protein in membrane preparations made from the cell line expressing the recombinant protein can be detected in HTS assays in one of several standard ways, including filtration of the protein-ligand complex to separate bound ligand from unbound ligand (Williams, Med. Res. Rev., 1991, 11, 147-184; Sweetnam et al., J. Natural Products, 1993, 56, 441-455). Alternative methods include a scintillation proximity assay (SPA) or a FlashPlate format in which such separation is unnecessary (Nakayama, Cur. Opinion Drug Disc. Dev., 1998, 1, 85-91 Bossé et al., J. Biomolecular Screening, 1998, 3, 285-292.). Binding of fluorescent ligands can be detected in various ways, including fluorescence energy transfer (FRET), direct spectrophotofluorometric analysis of bound ligand, or fluorescence polarization (Rogers, Drug Discovery Today, 1997, 2, 156-160; Hill, Cur. Opinion Drug Disc. Dev., 1998, 1, 92-97).

[00204] Other assays may be used to identify specific partners of a Rel_{Mtb}, including assays that identify partners of the target protein through measuring direct binding of test partners to the target protein, as well as assays that identify partners of target proteins through affinity ultrafiltration with ion spray mass spectroscopy/HPLC methods or other physical and analytical methods. Alternatively, such binding interactions are evaluated indirectly using the yeast two-hybrid system described in Fields et al., Nature, 340:245-246 (1989), and Fields et al., Trends in Genetics, 10:286-292 (1994), both of which are incorporated herein by reference. The two-hybrid system is a genetic assay for detecting interactions between two proteins or polypeptides. It can be used to identify proteins that bind to a known protein of interest, or to delineate domains or residues critical for an interaction. Variations on this methodology have been developed to clone genes that encode DNA binding proteins, to identify peptides that bind to a protein, and to screen for drugs. The two-hybrid system exploits the ability of a pair of

interacting proteins to bring a transcription activation domain into close proximity with a DNA binding domain that binds to an upstream activation sequence (UAS) of a reporter gene, and is generally performed in yeast. The assay requires the construction of two hybrid genes encoding (1) a DNA-binding domain that is fused to a first protein and (2) an activation domain fused to a second protein. The DNA-binding domain targets the first hybrid protein to the UAS of the reporter gene; however, because most proteins lack an activation domain, this DNA-binding hybrid protein does not activate transcription of the reporter gene. The second hybrid protein, which contains the activation domain, cannot by itself activate expression of the reporter gene because it does not bind the UAS. However, when both hybrid proteins are present, the noncovalent interaction of the first and second proteins tethers the activation domain to the UAS, activating transcription of the reporter gene. For example, when the first protein is a Rel gene product, or fragment thereof, that is known to interact with another protein or nucleic acid, this assay can be used to detect agents that interfere with the binding interaction. Expression of the reporter gene is monitored as different test agents are added to the system. The presence of an inhibitory agent results in lack of a reporter signal.

[00205] The yeast two-hybrid assay can also be used to identify proteins that bind to the gene product. In an assay to identify proteins that bind to a Rel_{Mtb} protein, or fragment thereof, a fusion polynucleotide encoding both a Rel_{Mtb} protein (or fragment) and a UAS binding domain (i.e., a first protein) may be used. In addition, a large number of hybrid genes each encoding a different second protein fused to an activation domain are produced and screened in the assay. Typically, the second protein is encoded by one or more members of a total cDNA or genomic DNA fusion library, with each second protein-coding region being fused to the activation domain. This system is applicable to a wide variety of proteins, and it is not even necessary to know the identity or function of the second binding protein. The system is highly sensitive and can detect interactions not revealed by other methods; even transient interactions may trigger transcription to produce a stable mRNA that can be repeatedly translated to yield the reporter protein.

[00206] Other assays may be used to search for agents that bind to the target protein. One such screening method to identify direct binding of test ligands to a target protein is described in U.S. Patent No. 5,585,277, incorporated herein by reference. This method relies on the principle that proteins generally exist as a mixture of folded and unfolded states, and continually alternate between the two states. When a test ligand binds to the folded form of a target protein (i.e., when the test ligand is a ligand of the target protein), the target protein molecule bound by the

ligand remains in its folded state. Thus, the folded target protein is present to a greater extent in the presence of a test ligand which binds the target protein, than in the absence of a ligand. Binding of the ligand to the target protein can be determined by any method that distinguishes between the folded and unfolded states of the target protein. The function of the target protein need not be known in order for this assay to be performed. Virtually any agent can be assessed by this method as a test ligand, including, but not limited to, metals, polypeptides, proteins, lipids, polysaccharides, polynucleotides and small organic molecules.

[00207] Another method for identifying ligands of a target protein is described in Wieboldt et al., Anal. Chem., 69:1683-1691 (1997), incorporated herein by reference. This technique screens combinatorial libraries of 20-30 agents at a time in solution phase for binding to the target protein. Agents that bind to the target protein are separated from other library components by simple membrane washing. The specifically selected molecules that are retained on the filter are subsequently liberated from the target protein and analyzed by HPLC and pneumatically assisted electrospray (ion spray) ionization mass spectroscopy. This procedure selects library components with the greatest affinity for the target protein, and is particularly useful for small molecule libraries.

[00208] Other embodiments of the invention comprise using competitive screening assays in which neutralizing antibodies capable of binding a polypeptide of the invention specifically compete with a test compound for binding to the polypeptide. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants with Rel_{Mtb}. Radiolabeled competitive binding studies are described in A.H. Lin et al. Antimicrobial Agents and Chemotherapy, 1997, vol. 41, no. 10. pp. 2127-2131, the disclosure of which is incorporated herein by reference in its entirety.

Identification of modulating agents

[00209] The invention also provides methods for identifying a modulator of binding between a Rel_{Mtb} and a Rel_{Mtb} binding partner, comprising the steps of: (a) contacting a Rel_{Mtb} binding partner and a composition comprising a Rel_{Mtb} in the presence and in the absence of a putative modulator compound; (b) detecting binding between the binding partner and the Rel_{Mtb}; and (c) identifying a putative modulator compound or a modulator compound in view of decreased or increased binding between the binding partner and the Rel_{Mtb} in the presence of the putative modulator, as compared to binding in the absence of the putative modulator. Following steps (a) and (b), compounds identified as modulating binding between Rel_{Mtb} and a Rel_{Mtb} binding

partner may be tested in other assays including, but not limited to, *in vivo* models, to confirm or quantitate modulation of binding to Rel_{Mtb}.

[00210] Rel_{Mtb} binding partners that stimulate Rel_{Mtb} activity are useful as agonists in disease states or conditions characterized by insufficient Rel_{Mtb} signaling (*e.g.*, as a result of insufficient activity of a Rel_{Mtb} ligand). Rel_{Mtb} binding partners that block ligand-mediated Rel_{Mtb} signaling are useful as Rel_{Mtb} antagonists to treat disease states or conditions characterized by excessive Rel_{Mtb} signaling. In addition Rel_{Mtb} modulators in general, as well as Rel_{Mtb} polynucleotides and polypeptides, are useful in diagnostic assays for such diseases or conditions.

[00211] In some aspects, the invention provides methods for treating a disease or abnormal condition by administering to a patient in need of such treatment a substance that modulates the activity or expression of a polypeptide having sequences of SEQ ID NOs:7-13.

[00212] Agents that modulate (i.e., increase, decrease, or block) Rel_{Mtb} activity or expression may be identified by incubating a putative modulator with a cell containing a Rel_{Mtb} polypeptide or polynucleotide and determining the effect of the putative modulator on Rel_{Mtb} activity or expression. The selectivity of a compound that modulates the activity of Rel_{Mtb} can be evaluated by comparing its effects on Rel_{Mtb} to its effect on other related proteins. Following identification of compounds that modulate Rel_{Mtb} activity or expression, such compounds may be further tested in other assays including, but not limited to, *in vivo* models, in order to confirm or quantitate their activity. Selective modulators may include, for example, antibodies and other proteins, peptides, or organic molecules that specifically bind to a Rel_{Mtb} polypeptide or a Rel_{Mtb}-encoding nucleic acid. Modulators of Rel_{Mtb} activity will be therapeutically useful in treatment of diseases and physiological conditions in which normal or aberrant Rel_{Mtb} activity is involved. Rel_{Mtb} polynucleotides, polypeptides, and modulators may be used in the treatment of such diseases and conditions as infections, such as tuberculosis, staph, strep, and other bacterial infections.

[00213] Methods of the invention to identify modulators include variations on any of the methods described above to identify binding partner compounds, the variations including techniques wherein a binding partner compound has been identified and the binding assay is carried out in the presence and absence of a candidate modulator. A modulator is identified in those instances where binding between the Rel_{Mtb} polypeptide and the binding partner compound changes in the presence of the candidate modulator compared to binding in the absence of the candidate modulator compound. A modulator that increases binding between the Rel_{Mtb} polypeptide and the binding partner compound is described as an enhancer or activator, and a modulator that decreases binding between the Rel_{Mtb} polypeptide and the binding partner

compound is described as an inhibitor. Following identification of modulators, such compounds may be further tested in other assays including, but not limited to, *in vivo* models, in order to confirm or quantitate their activity as modulators.

[00214] The invention also comprehends high-throughput screening (HTS) assays to identify compounds that interact with or inhibit biological activity (i.e., affect enzymatic activity, binding activity, etc.) of a Rel_{Mtb} polypeptide. HTS assays permit screening of large numbers of compounds in an efficient manner. Cell-based HTS systems are contemplated to investigate Rel_{Mtb} protein-ligand interaction. HTS assays are designed to identify "hits" or "lead compounds" having the desired property, from which modifications can be designed to improve the desired property. Chemical modification of the "hit" or "lead compound" is often based on an identifiable structure/activity relationship between the "hit" and the Rel_{Mtb} polypeptide.

[00215] Another aspect of the present invention is directed to methods of identifying compounds which modulate (i.e., increase or decrease) an activity of Rel_{Mtb} comprising contacting Rel_{Mtb} with a compound, and determining whether the compound modifies activity of Rel_{Mtb}. The activity in the presence of the test compared is measured to the activity in the absence of the test compound. Where the activity of the sample containing the test compound is higher than the activity in the sample lacking the test compound, the compound will have increased activity. Similarly, where the activity of the sample containing the test compound is lower than the activity in the sample lacking the test compound, the compound will have inhibited activity. Following the identification of compounds that modulate an activity of Rel_{Mtb}, such compounds can be further tested in other assays including, but not limited to, *in vivo* models, in order to confirm or quantitate their activity.

[00216] The present invention is particularly useful for screening compounds by using Rel_{Mtb} in any of a variety of drug screening techniques. The compounds to be screened include (which may include compounds which are suspected to modulate Rel_{Mtb} activity), but are not limited to, extracellular, intracellular, biologic or chemical origin. The Rel_{Mtb} polypeptide employed in such a test may be in any form, preferably, free in solution, attached to a solid support, borne on a cell surface or located intracellularly. One skilled in the art can, for example, measure the formation of complexes between Rel_{Mtb} and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between Rel_{Mtb} and its substrate caused by the compound being tested. In some embodiments, fragments comprising the domains of Rel_{Mtb} are used.

[00217] The activity of Rel_{Mtb} polypeptides of the invention can be determined by, for example, examining the ability to hydrolyze or synthesize (p)ppGpp.

[00218] The modulators of the invention exhibit a variety of chemical structures, which can be generally grouped into non-peptide mimetics of natural Rel_{Mtb} ligands, peptide and non-peptide allosteric effectors of Rel protein, and peptides that may function as activators or inhibitors (competitive, uncompetitive and non-competitive) (*e.g.*, antibody products) of Rel_{Mtb} proteins. The invention does not restrict the sources for suitable modulators, which may be obtained from natural sources such as plant, animal or mineral extracts, or non-natural sources such as small molecule libraries, including the products of combinatorial chemical approaches to library construction, and peptide libraries.

[00219] Other assays can be used to examine enzymatic activity including, but not limited to, photometric, radiometric, HPLC, electrochemical, and the like, which are described in, for example, Enzyme Assays: A Practical Approach, eds. R. Eisenthal and M. J. Danson, 1992, Oxford University Press, which is incorporated herein by reference in its entirety.

[00220] The use of cDNAs encoding proteins in drug discovery programs is well-known; assays capable of testing thousands of unknown compounds per day in high-throughput screens (HTSs) are thoroughly documented. The literature is replete with examples of the use of radiolabeled ligands in HTS binding assays for drug discovery (see Williams, Medicinal Research Reviews, 1991, 11, 147-184.; Sweetnam, et al., J. Natural Products, 1993, 56, 441-455 for review). Recombinant proteins are preferred for binding assay HTS because they allow for better specificity (higher relative purity), provide the ability to generate large amounts of protein material, and can be used in a broad variety of formats (see Hodgson, Bio/Technology, 1992, 10, 973-980; each of which is incorporated herein by reference in its entirety).

[00221] A variety of heterologous systems is available for functional expression of recombinant proteins that are well known to those skilled in the art. Such systems include bacteria (Strosberg, et al., Trends in Pharmacological Sciences, 1992, 13, 95-98), yeast (Pausch, Trends in Biotechnology, 1997, 15, 487-494), several kinds of insect cells (Vanden Broeck, Int. Rev. Cytology, 1996, 164, 189-268), amphibian cells (Jayawickreme et al., Current Opinion in Biotechnology, 1997, 8, 629-634) and several mammalian cell lines (CHO, HEK-293, COS, etc.; see Gerhardt, et al., Eur. J. Pharmacology, 1997, 334, 1-23). These examples do not preclude the use of other possible cell expression systems, including cell lines obtained from nematodes (PCT application WO 98/37177).

[00222] In some embodiments of the invention, methods of screening for compounds that modulate Rel_{Mtb} activity comprise contacting test compounds with Rel_{Mtb} and assaying for the presence of a complex between the compound and Rel_{Mtb}. In such assays, the ligand is typically labeled. After suitable incubation, free ligand is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular compound to bind to Rel_{Mtb}.

[00223] Candidate modulators contemplated by the invention include compounds selected from libraries of either potential activators or potential inhibitors. There are a number of different libraries used for the identification of small molecule modulators, including: (1) chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules. Chemical libraries consist of random chemical structures, some of which are analogs of known compounds or analogs of compounds that have been identified as "hits" or "leads" in other drug discovery screens, some of which are derived from natural products, and some of which arise from non-directed synthetic organic chemistry. Natural product libraries are collections of microorganisms, animals, plants, or marine organisms which are used to create mixtures for screening by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of plants or marine organisms. Natural product libraries include polyketides, non-ribosomal peptides, and variants (non-naturally occurring) thereof. For a review, see Science 282:63-68 (1998). Combinatorial libraries are composed of large numbers of peptides, oligonucleotides, or organic compounds as a mixture. These libraries are relatively easy to prepare by traditional automated synthesis methods, PCR, cloning, or proprietary synthetic methods. Of particular interest are non-peptide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to modulate activity.

[00224] Still other candidate inhibitors contemplated by the invention can be designed and include soluble forms of binding partners, as well as such binding partners as chimeric, or fusion, proteins. A "binding partner" as used herein broadly encompasses non-peptide modulators, as well as peptide modulators other than natural ligands, antibodies, antibody fragments, and

modified compounds comprising antibody domains that are immunospecific for the expression product of the identified Rel_{Mtb} gene.

[00225] The polypeptides of the invention are employed as a research tool for identification, characterization and purification of interacting, regulatory proteins. Appropriate labels are incorporated into the polypeptides of the invention by various methods known in the art and the polypeptides are used to capture interacting molecules. For example, molecules are incubated with the labeled polypeptides, washed to remove unbound polypeptides, and the polypeptide complex is quantified. Data obtained using different concentrations of polypeptide are used to calculate values for the number, affinity, and association of polypeptide with the protein complex.

[00226] Labeled polypeptides are also useful as reagents for the purification of molecules with which the polypeptide interacts including, but not limited to, inhibitors. In one embodiment of affinity purification, a polypeptide is covalently coupled to a chromatography column. Cells and their membranes are extracted, and various cellular subcomponents are passed over the column. Molecules bind to the column by virtue of their affinity to the polypeptide. The polypeptide-complex is recovered from the column, dissociated and the recovered molecule is subjected to protein sequencing. This amino acid sequence is then used to identify the captured molecule or to design degenerate oligonucleotides for cloning the corresponding gene from an appropriate cDNA library.

[00227] Alternatively, compounds may be identified which exhibit similar properties to the ligand for the Rel_{Mtb} of the invention, but which are smaller and exhibit a longer half time than the endogenous ligand in a human or animal body. When an organic compound is designed, a molecule according to the invention is used as a "lead" compound. The design of mimetics to known pharmaceutically active compounds is a well-known approach in the development of pharmaceuticals based on such "lead" compounds. Mimetic design, synthesis and testing are generally used to avoid randomly screening a large number of molecules for a target property. Furthermore, structural data deriving from the analysis of the deduced amino acid sequences encoded by the DNAs of the present invention are useful to design new drugs, more specific and therefore with a higher pharmacological potency.

[00228] In a particular embodiment, the novel molecules identified by the screening methods according to the invention are low molecular weight organic molecules, in which case a composition or pharmaceutical composition can be prepared thereof for oral intake, such as in tablets. The compositions, or pharmaceutical compositions, comprising the nucleic acid

molecules, vectors, polypeptides, antibodies and compounds identified by the screening methods described herein, can be prepared for any route of administration including, but not limited to, oral, intravenous, cutaneous, subcutaneous, nasal, intramuscular or intraperitoneal. The nature of the carrier or other ingredients will depend on the specific route of administration and particular embodiment of the invention to be administered. Examples of techniques and protocols that are useful in this context are, inter alia, found in Remington's Pharmaceutical Sciences, 16th edition, Osol, A (ed.), 1980, which is incorporated herein by reference in its entirety.

[00229] The dosage of these low molecular weight compounds will depend on the disease state or condition to be treated and other clinical factors such as weight and condition of the human or animal and the route of administration of the compound. For treating human or animals, between approximately 0.5 mg/kg of body weight to 500 mg/kg of body weight of the compound can be administered. Therapy is typically administered at lower dosages and is continued until the desired therapeutic outcome is observed.

[00230] The present compounds and methods, including nucleic acid molecules, polypeptides, antibodies, compounds identified by the screening methods described herein, have a variety of pharmaceutical applications and may be used, for example, to treat or prevent unregulated cellular growth, such as cancer cell and tumor growth. In a particular embodiment, the present molecules are used in gene therapy. For a review of gene therapy procedures, see *e.g.* Anderson, Science, 1992, 256, 808-813, which is incorporated herein by reference in its entirety.

[00231] The present invention also encompasses a method of agonizing (stimulating) or antagonizing a Rel_{Mtb} natural binding partner associated activity in a mammal comprising administering to said mammal an agonist or antagonist to one of the above disclosed polypeptides in an amount sufficient to effect said agonism or antagonism. One embodiment of the present invention, then, is a method of treating diseases in a mammal with an agonist or antagonist of the protein of the present invention comprises administering the agonist or antagonist to a mammal in an amount sufficient to agonize or antagonize Rel_{Mtb}-associated Function: S.

[00232] Methods of determining the dosages of compounds to be administered to a patient and modes of administering compounds to an organism are disclosed in International patent publication number WO 96/22976, published August 1 1996, which is incorporated herein by reference in its entirety, including any drawings, figures or tables. Those skilled in the art will

appreciate that such descriptions are applicable to the present invention and can be easily adapted to it.

[00233] The proper dosage depends on various factors such as the type of disease being treated, the particular composition being used and the size and physiological condition of the patient. Therapeutically effective doses for the compounds described herein can be estimated initially from cell culture and animal models. For example, a dose can be formulated in animal models to achieve a circulating concentration range that initially takes into account the IC50 as determined in cell culture assays. The animal model data can be used to more accurately determine useful doses in humans.

[00234] Plasma half-life and biodistribution of the drug and metabolites in the plasma, tumors and major organs can also be determined to facilitate the selection of drugs most appropriate to inhibit a disorder. Such measurements can be carried out. For example, HPLC analysis can be performed on the plasma of animals treated with the drug and the location of radiolabeled compounds can be determined using detection methods such as X-ray, CAT scan and MRI. Compounds that show potent inhibitory activity in the screening assays, but have poor pharmacokinetic characteristics, can be optimized by altering the chemical structure and retesting. In this regard, compounds displaying good pharmacokinetic characteristics can be used as a model.

[00235] Toxicity studies can also be carried out by measuring the blood cell composition. For example, toxicity studies can be carried out in a suitable animal model as follows: 1) the compound is administered to mice (an untreated control mouse should also be used); 2) blood samples are periodically obtained via the tail vein from one mouse in each treatment group; and 3) the samples are analyzed for red and white blood cell counts, blood cell composition and the percent of lymphocytes versus polymorphonuclear cells. A comparison of results for each dosing regime with the controls indicates if toxicity is present.

[00236] At the termination of each toxicity study, further studies can be carried out by sacrificing the animals (preferably, in accordance with the American Veterinary Medical Association guidelines Report of the American Veterinary Medical Assoc. Panel on Euthanasia, Journal of American Veterinary Medical Assoc., 202:229-249, 1993). Representative animals from each treatment group can then be examined by gross necropsy for immediate evidence of metastasis, unusual illness or toxicity. Gross abnormalities in tissue are noted and tissues are examined histologically. Compounds causing a reduction in body weight or blood components

are less preferred, as are compounds having an adverse effect on major organs. In general, the greater the adverse effect the less preferred the compound.

[00237] For the treatment of many diseases, the expected daily dose of a hydrophobic pharmaceutical agent is between 1 to 500 mg/day, preferably 1 to 250 mg/day, and most preferably 1 to 50 mg/day. Drugs can be delivered less frequently provided plasma levels of the active moiety are sufficient to maintain therapeutic effectiveness. Plasma levels should reflect the potency of the drug. Generally, the more potent the compound the lower the plasma levels necessary to achieve efficacy.

[00238] The attached Sequence Listing contains the sequences of the polynucleotides and polypeptides of the invention and is incorporated herein by reference in its entirety.

[00239] The invention is now described with reference to the following examples. These examples are provided for the purpose of illustration only and the invention should in no way be construed as being limited to these examples, but rather should be construed to encompass any and all variations which become evident as a result of the teaching provided herein. Those of skill in the art will readily recognize a variety of non-critical parameters that could be changed or modified to yield essentially similar results. The disclosures of each and every patent, patent application, and publication cited herein are hereby incorporated herein by reference in their entirety. While this invention has been disclosed with reference to specific embodiments, it is apparent that other embodiments and variations of this invention may be devised by others skilled in the art without departing from the true spirit and scope of the invention. The appended claims are intended to be construed to include all such embodiments and equivalent variations.

EXAMPLES

Example 1. DNA manipulation and PCR amplification.

[00240] Rel_{Mtb} fragments were PCR-amplified from full-length Rel_{Mtb} template DNA using Vent Polymerase (New England Biolabs). The following primers were used: N-term, GATATA*CCATGG*GCAGCAGCC (SEQ ID NO:1); 87-394 N-term, GATTGATC*CATATG*GGTTACACCCTGGAGGCGTT (SEQ ID NO:2); 1-394 and 87-394 C-term, GATCAT*GGATCC*CTAGTCGTAGCGCAATGATTCCA (SEQ ID NO:3); 1-203 C-term, GATCAT*GGATCC*CTACTCCTCGTACTTCTTGGGAT (SEQ ID NO:4); 1-181 C-term, GATCAT*GGATCC*CTACATGCCCAGCCGATGCGCCAG (SEQ ID NO:5); 1-156 C-term, GATCAT*GGATCC*CTACAAGAAGCGCATGGTGCGCAT (SEQ ID NO:6). Restriction sites

are italicized. Amplification conditions were 94 °C 45s, 55 °C 45s, 72 °C 3 min for 30 cycles. The PCR products were digested with Nde1, NcoI, BamH1 (New England Biolabs) and cloned into pET15b vectors.

Example 2. Expression And Purification Of RelMtb.

[00241] BL21 [DE3] cells carrying pET15b plasmids with the appropriate fragment DNA were grown in LB media to an OD of 0.6 for 37 °C growth (1-394, 1-450, 1-181) and 1.1 for 15 °C growth (1-203, 87-394). Isopropyl-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1.0 mM (37 °C growth) and 0.04 mM (15 °C growth) and the protein was allowed to express for 3 hours. In the case of 15 °C growth, the cell solution was allowed to cool from 37 °C to 30 °C from an O.D. of 0.9-1.1 and then cooled further from 30 °C to 15 °C for two hours prior to induction with 0.04 mM IPTG. The solution was then allowed to express for 15 hours. Cells were resuspended in buffer containing 50 mM NaPi (pH 8.0), 300 mM NaCl, and 20mM imidazole and frozen at –80 °C until purification. Cells were broken by 3 rounds of thawing and freezing, followed by addition of 20 mg/L of lysozyme (Boehringer Mannheim). A pellet of protease inhibitor cocktail (Complete™, EDTA-Free, Boehringer-Mannheim) was added, and the solution was allowed to sit for 20 minutes, until viscosity was observed. 10 Units of DNase were then added, and the solution was sonicated at 75mHz for 5 x 30 seconds. Cell debris was removed by two rounds of centrifugation. The lysate was added to Ni2+ -NTA agarose (Qiagen) for 1 hr and the resin was then washed with a buffer of 50 mM NaPi (pH 8.0), 300 mM NaCl, and 20 mM imidazole for a further 20 minutes. The resin was then transferred to a column and again washed with 50 mL of 50 mM NaPi (pH 8.0), 300 mM NaCl, and 20 mM imidazole. Protein was eluted with buffers containing 50 mM NaPi (pH 8.0), 300 mM NaCl, 50 mM imidazole (10 1.5 mL fractions), 100 mM imidazole (10 1 mL fractions) and 250 mM imidazole (6-10 0.5 mL fractions). Protein fractions were checked at OD280 using calculated extinction coefficients .296 M-1cm-1 for 1-156 (17.3 kDa) fragment, .255 M-1cm-1 for 1-181 (20.1 kDa) fragment, .533 M-1cm-1 for 1-203 (22.7 kDa) fragment, 1.412 M-1cm-1 for 87-394 (35.1 kDa) fragment, and 1.199 M-1cm-1 for 1-394 (44.6 kDa) fragment (Swiss-Prot Analysis), and purity was assessed by SDS-PAGE analysis. Pure fractions were then concentrated via centrifugal filter (Millipore) and assayed for activity.

[00242] The pET22b expression system (Novagen) was used to purify the 82 kDa recombinant Rel_{Mtb} protein as previously described (5). Concentration of wild type Rel_{Mtb} was determined by A280 (Beckman DU 640) using a calculated extinction coefficient of 0.979 M-1cm-1 for full-length 82 kDa Rel_{Mtb} (Swiss-Prot Analysis).

Example 3. Transferase Assays

[00243] Ribosome-independent transferase reactions (30 °C) contained 50 mM HEPES (pH 8.0), 100-225 mM NaCl, 1 mM DTT, indicated concentrations of ATP/GTP (Pharmacia), either 200 nM 738 amino acid full-length Rel_{Mtb}, 1.44 μM 1-394, or 1.3 μM 87-394, and varying [MgCl2] or [MnCl2] and were carried out as described previously (6). In addition, the Rel_{Mtb} activating complex reactions included ribosomes (0.15-0.30 μM), tRNA (0.20-1.50 μM), and mRNA (2.00 μM; Poly(A) and Poly(U) from Boehringer-Mannheim, Poly(AU) from Sigma). After the reaction components were mixed, final pH was determined using an NMR pH electrode (Wilmad).

[00244] Assay method 1: Reactions were monitored using either [γ-33P]ATP or [γ-33P]GTP (NEN) at 1μCi/μmol. Reaction rates were calculated by taking 5 μl aliquots at multiple time points, spotting on PEI-cellulose TLC plates (Sigma-Aldrich) and developing in 1.5 M KPi (pH 3.4). Reaction products were visualized using a Storm Phosphorimager (Molecular Dynamics) with an enhanced sensitivity storage phosphor screen, and images quantitated using ImageQuant version 1.2 (Molecular Dynamics). All reaction rates were linear with both time and enzyme concentration under these conditions, and less than 10% of substrate was converted to product.

[00245] Assay for (p)ppGpp Hydrolysis: Hydrolysis reactions (30 °C) contained 50 mM HEPES (pH 8.0), 100-225 mM NaCl, 1 mM DTT, indicated [(p)ppGpp], either 200 nM 738 amino acid full-length Rel_{Mtb}, 1.44 μM 1-394, 2.4 μM 1-203, 52.1 μM 1-181, or 52.1 μM 1-156, and varying MnCl2 concentrations and were carried out as described previously (6). Preparative amounts of pppGpp were synthesized from partially purified purine nucleotide pyrophosphotransferase (PPK) from S. morookaensis. S. morookaensis PPK was prepared as previously described (5). (p)ppGpp* was prepared in a reaction containing 100 mM glycine (pH 10.0), 20 mM MgCl2, [γ-33P]ATP (15 mM, 1 μCi/μmol), 20 mM GTP, 25 μl of PPK for 40 minutes at 37 °C. At the end of the incubation, the reaction was brought to pH 7.0 using 1 N HCl, and (p)ppGpp was isolated using an ÄKTA FPLC with a High Trap anion exchange column. (p)ppGpp was dialyzed (100 Da MWCO, Spectrum) to remove NaCl and lyophilized. NaCl concentration in the (p)ppGpp preparation was measured using a Radiometer CDM210 conductivity meter. Concentrations of (p)ppGpp were determined spectrophotometrically using ε252=1.37 X 104 M-1 cm-1; purity was confirmed by running on TLC with other nucleotide standards.

[00246] Assay 1: The hydrolysis reaction product, *PPi, was visualized and quantitated as in Assay (1) for the transferase reaction.

[00247] Assay 2: PPi was quantitated using a continuous spectrophotometric assay by following the decrease in A340 when two moles of NADH are oxidized to NAD+ per mole of pyrophosphate consumed (8). Reactions were set up and carried out as described previously (6).

Example 4. Ribosomes, tRNA Synthetase, tRNAs

[00248] Ribosomes, tRNA Synthetase, tRNAs were prepared as previously described (6).

Example 5. Gel Filtration Chromatography.

[00249] A Superdex 200 HR 10/30 column (24 mL bed volume) was used to estimate the apparent molecular weight of Rel_{Mtb} multimers at room temperature. Protein samples were applied to the column system equilibrated in a standard elution buffer (50 mM HEPES buffer (pH = 8.2) and 200 mM NaCl) at a flow rate of .75 mL/min. Ribonuclease A, Chymotrypsinogen A, Ovalbumin, Albumin, Aldolase, Catalase, Ferritin, and Thyroglobulin (Amersham Pharmacia Biotech) were used as protein standards for column calibration. A Kav vs log molecular weight calibration curve was prepared to estimate the apparent molecular weight of the enzyme.

$$Kav = \frac{Ve - Vo}{Vt - Vo}$$

Where Ve = elution volume for the protein

Vo = column void volume = elution volume for Blue Dextran 2000

Vt = total bed volume

Elution was monitored for protein at A280.

Example 6. Dissection of Bifunctional Protein into Separate Catalytic Domains.

[00250] To investigate if the full-length 738 amino acid Rel_{Mtb} can both synthesize and hydrolyze (p)ppGpp we characterized fragments of full-length Rel_{Mtb} to learn more about the domains of the protein as well as provide fragments for further biophysical studies. Six fragments were purified and four fragments were characterized in detail: (i) 1-394 amino acid fragment capable of both (p)ppGpp synthesis and hydrolysis (ii) 1-181 amino acid fragment capable of only (p)ppGpp hydrolysis (iii) 1-203 amino acid fragment capable of only (p)ppGpp hydrolysis (iv) and an 87-394 amino acid fragment capable of only (p)ppGpp synthesis (Figure 1).

Characterization and Comparison of Dual-Function RelMtb Fragments to Wild Type RelMtb.

[00251] In order to determine whether both of the catalytic activities of full-length Rel_{Mtb} could be localized to a smaller region, two truncated proteins were purified: a 1-450 and a 1-394 amino acid fragment (Figure 1). Both fragments were capable of synthesis and hydrolysis. The 1-394 amino acid fragment had a kcat for (p)ppGpp synthesis about 3-fold lower than that of wild type Rel_{Mtb} (Table 4). The 1-394 amino acid fragment Km for GTP decreased about 1.2-fold and the Km for ATP increased almost 1.5-fold resulting in slightly lower kcat/Km values relative to the full-length protein (Table 4). The kcat for (p)ppGpp hydrolysis decreased 70-fold while the Km for (p)ppGpp decreased 1.3-fold resulting in a 55-fold decrease in kcat/Km relative to wild type Rel_{Mtb} (Table 5). Full-length Rel_{Mtb} kinetic numbers are used from our last paper (6), however the full-length Rel_{Mtb} was assayed in each experiment and used as a control.

Table 4: Kinetic constants for RelMtb transferase reaction

Reaction	K_{ATP} (mM)	K_{GTP} (mM)	k_{cat} (s^{-1})	k_{cat}/K_{ATP} ($mM^{-1}\ s^{-1}$)	k_{cat}/K_{GTP} ($mM^{-1}\ s^{-1}$)
1) Full-length Ribosome-Independent Transferase (Mg^{2+})	1.96 ± 0.24	1.38 ± 0.16	1.21 ± 0.16	0.62	0.88
2)Full-length Rel_{Mtb} + Uncharged tRNA	1.08 ± 0.13	0.73 ± 0.12	1.28 ± 0.18	1.19	1.75
3) Rel_{Mtb} + Uncharged tRNA• Ribosomes•mRNA (RAC)	0.45 ± 0.08	0.31 ± 0.06	24.68 ± 2.87	54.84	79.61
4) 1-394 Ribosome-Independent Transferase (Mg^{2+})	2.83 ± 0.44	1.12 ± 0.15	0.24 ± 0.03	0.08	0.21
5) 87-394 Ribosome-Independent Transferase (Mg^{2+})	2.47 ± 0.56	1.11 ± 0.16	0.18 ± 0.06	0.07	0.16
6) 1-394 + Uncharged tRNA	2.81 ± 0.45	1.01 ± 0.08	0.25 ± 0.04	0.09	0.25
7) 1-394 + Uncharged tRNA• Ribosomes•mRNA (RAC)	2.75 ± 0.41	1.14 ± 0.08	0.19 ± 0.06	0.07	0.17
8) 1-394 monomer Ribosome-Independent	NA	1.31 ± 0.18	0.52 ± 0.04	NA	0.40
9) 1-394 trimer Ribosome-Independent	NA	1.29 ± 0.33	0.16 ± 0.02	NA	0.12

Table 4: a The transferase reaction assayed was: *p-p-p-A (*ATP) + p-p-p-G (GTP) ↔ p-A (AMP) + p-p-p-G-p-p*.. p-p-p-G-p-p* indicates that the radioactive label is on the 3'-PPi of the ribose ring of GTP. Reactions were performed at 30°C, pH 8.0. For KATP determination, GTP was held constant at saturating levels (15-20 mM, 1 μCi/μMol [γ-33P]GTP); for KGTP determination, ATP was held constant (15-20 mM, 1 μCi/μMol [γ-33P]ATP). The kcat was calculated as if all components in reaction were monomers (1 mole trimer (135 kDa) equals 3 moles monomer (45 kDa) .

Table 5: Kinetic constants for Rel_{Mtb} hydrolysis reaction

Reaction	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1}\ s^{-1}$)

1) Full-length Rel_{Mtb} + pppGpp (Basal level)	0.48 ± 0.06	2.83 ± 0.33	5.90
2) 1-394 Rel_{Mtb} + pppGpp	0.37 ± 0.09	0.040 ± 0.006	0.108
3) 1-203 Rel_{Mtb} + pppGpp	0.43 ± 0.10	0.033 ± 0.006	0.077
4) 1-181 Rel_{Mtb} + pppGpp	0.51 ± 0.11	0.0043 ± 0.0006	0.0084

Table 5: a The hydrolysis reactions assayed are: (1) p-p-p-G-p-p* (pppGpp) → *PPi + p-p-p-G (GTP); p-p-p-G-p-p* indicates that the radioactive label is on the 3'-PPi of the ribose ring of GTP. Reactions were performed at 30°C, pH 8.0.

Characterization and Comparison of (p)ppGpp Hydrolysis Only Fragments to Wild Type Rel_{Mtb}.

[00252] Three fragments were purified and assayed to determine the domain necessary for (p)ppGpp hydrolysis: 1-203, 1-181, and 1-156 amino acid fragments (Figure 1). Both the 1-203 and 1-181 amino acid fragments possessed (p)ppGpp hydrolysis activity but were not capable of (p)ppGpp synthesis activity. The 1-156 amino acid fragment was not capable of either (p)ppGpp synthesis or hydrolysis. The kcat for the 1-203 amino acid fragment decreased 85-fold compared to wild type Rel_{Mtb} while the Km for (p)ppGpp was approximately the same, resulting in a nearly 80-fold decrease in kcat/Km for (p)ppGpp hydrolysis (Table 5). The kcat for the 1-181 fragment was decreased almost 700-fold compared to wild type Rel_{Mtb} while the Km for (p)ppGpp was approximately the same, resulting in a 700-fold decrease in kcat/Km for (p)ppGpp hydrolysis (Table 5). Once again, full-length Rel_{Mtb} kinetic numbers are used from our last paper (6), however the full-length Rel_{Mtb} was assayed in each experiment and used as a control.

Characterization and Comparison of a (p)ppGpp Synthesis Only Fragment to Wild Type RelMtb.

[00253] A fragment containing amino acids 87-394 was constructed and found to have only (p)ppGpp synthesis activity. The kcat for the 87-394 protein fragment was 2.5-fold lower than that of wild type Rel_{Mtb}, while the Km for ATP and GTP was similar to that of wild type Rel_{Mtb} (Table 4).

Effects of Divalent Cations on Transferase and Hydrolysis Rates.

[00254] The optimal Mg2+ concentration for the wild type catalyzed (p)ppGpp synthesis reaction was approximately equal to the total concentration of nucleotide substrates, [ATP + GTP], whereas the optimal Mn2+ concentration was approximately one-half the total concentration of nucleotide substrates. The 1-394 and 87-394 Rel_{Mtb} fragments also had the same metal concentration range requirements for the transferase reaction as wild type Rel_{Mtb}

(Figure 2A). Adding divalent cations in excess of [ATP + GTP] decreased (p)ppGpp synthesis in a manner similar to what we reported for wild type Rel_{Mtb} (6).

[00255] Similar to wild type Rel_{Mtb}, only Mn2+ is capable of supporting the hydrolysis reaction for the truncated Rel_{Mtb} fragments. The optimal Mn2+ concentration for the wild type Rel_{Mtb}-catalyzed hydrolysis reaction was a 2-3 fold molar excess over the nucleotide substrate, (p)ppGpp. This requirement contrasts with that of the 1-181, 1-203, and 1-394 amino acid fragments, where a 6-10 fold molar excess of Mn2+ over (p)ppGpp substrate was the optimal concentration (Figure 3 A,B,C). Although a broad range of Mn2+ at concentrations well above substrate levels supported an optimal hydrolysis activity, Rel_{Mtb} hydrolysis activity did eventually decrease with very high levels of Mn2+.

Higher-Level Regulation of RelMtb Activity by Uncharged tRNA and the RelMtb Activating Complex (RAC).

[00256] In our previous study, adding only uncharged tRNA to the full-length Rel_{Mtb} ribosome-independent transferase reaction caused a 2-fold decrease in KATP and KGTP (6). Addition of a complex of uncharged tRNA, ribosomes, and cognate mRNA (RAC) to the wild type Rel_{Mtb} catalyzed transferase reaction lowered the KATP and KGTP 4-fold and increased kcat 20-fold. However, our present characterization of the Rel_{Mtb} fragments with (p)ppGpp transferase activity (1-394, 87-394) indicates that the transferase activity fails to be activated above basal levels by uncharged tRNA alone or the Rel_{Mtb} activating complex (RAC) (Table 4).

[00257] Also in our previous study, the addition of uncharged tRNA or the RAC to wild type Rel_{Mtb} had the opposite effect on hydrolysis activity, increasing K(p)ppGpp 3-fold, and decreasing kcat 2-fold. In the present study, the uncharged tRNA or the RAC failed to inhibit the activity of the fragments capable of (p)ppGpp hydrolysis (1-181, 1-203, 1-394).

Evidence of RelMtb Multimerization.

[00258] To determine whether multimerization may play a role in the regulation of Rel_{Mtb}, the protein was studied using analytical gel filtration. Wild type Rel_{Mtb} formed one distinct peak with an apparent molecular mass of 240 kDa that corresponded to a trimer state (Figure 4A). The 1-394 amino acid fragment formed two distinct peaks with apparent molecular weights of 141 kDa and 47 kDa, which corresponded to a trimer and monomer state, respectively (Figure 4B). The 1-181 amino acid fragment formed one distinct peak with an apparent molecular mass of 27.8 kDa (calculated molecular mass of 20 kDa), which is consistent with a monomer state (Figure 4C). Similarly the 1-203 fragment formed one distinct peak with an apparent molecular

mass of 31 kDa, (calculated molecular mass of 23 kDa) which approximately corresponds to the molecular mass of the monomer state (Figure 4D).

Stability of RelMtb Multimers.

[00259] Incubation of wild type Rel_{Mtb} with 1mM DTT followed by running on the HPLC in standard elution buffer with 1mM DTT did not disrupt the stability of the trimer. Incubation of the 1-394 Rel_{Mtb} fragment with 1mM DTT followed by running on the HPLC in elution buffer with 1mM DTT did not affect the stability of the multimer peak. The monomer and trimer peaks of the 1-394 fragment were each isolated (0.14 mg/mL and 0.073 mg/mL, respectively) and injected back onto the HPLC column after sitting on ice for both 30 minutes and 24 hours. Under these conditions there was no interconversion between monomer and trimer. The 1-394 trimer and monomer were run on denaturing SDS PAGE and both formed single bands corresponding to a 44 kDa fragment, the size of the monomer 1-394 fragment.

Activity of Isolated 1-394 Monomer and Trimer.

[00260] After separating the 1-394 monomer from the 1-394 trimer on the HPLC, the isolated monomer and isolated trimer were each assayed for (p)ppGpp synthesis and hydrolysis activity. Reaction mixtures contained equal concentrations of total protein (0.073 mg/mL), or 1.62 μM 1-394 monomer (45 kDa) for monomer reaction and .541 μM 1-394 trimer (135 kDa) for trimer reaction. The (p)ppGpp synthesis kcat of the isolated monomers was three times that of the isolated trimers at the same concentration (Table 4). The synthesis kcat of the 1-394 fragment (mixture of monomers and trimers) at the same concentration was between that of the monomer and trimer (Table 4). The isolated monomers and trimers still had a (p)ppGpp synthesis kcat lower than the basal synthesis kcat of wild type Rel_{Mtb} at the same concentration (Table 4). The (p)ppGpp hydrolysis rate of the isolated monomers and trimers at the same concentration was similar to each other and similar to the rate of the 1-394 fragment (mixture of monomers and trimers) at the same concentration (data not shown). The isolated monomers and trimers still had a (p)ppGpp hydrolysis rate lower (45-fold and 50-fold respectively) than the basal synthesis rate of wild type Rel_{Mtb} at the same concentration.

Implementation:

[00261] In the gram-negative bacteria E. coli two proteins are responsible for the synthesis and hydrolysis of (p)ppGpp: the RelA protein synthesizes (p)ppGpp in response to amino acid starvation and the SpoT protein hydrolyzes (p)ppGpp. SpoT is also capable of synthesizing (p)ppGpp, but RelA is not capable of hydrolyzing (p)ppGpp. Results of deletion analysis (10) of the spoT gene in E. coli found mutants that contain either one or the other activity: residues 1-

203 contain the domain for (p)ppGpp degradation while residues 85-375 contain the domain for (p)ppGpp synthesis. These results are consistent with those found in Rel_{Mtb} (Figure 5E,G).

[00262] The SpoT protein is related to the gram-positive, dual-function Rel class of proteins, and it is theorized that after gene duplication of an ancestral rel-like gene, the spoT and relA genes evolved from these duplicated genes (11). Thus, it is likely that Rel_{Mtb} from the gram-positive organism Mycobacterium tuberculosis, which carries out both synthesis and hydrolysis of (p)ppGpp, will have separate catalytic domains like the SpoT protein of E. coli (10). While this manuscript was in preparation, Mechold and colleagues also demonstrated separate catalytic domains in a Rel homolog in Streptomyces equisimilis: residues 1-224 contain the domain for (p)ppGpp degradation and residues 79-385 contain the domain for (p)ppGpp synthesis (12). However, in our study Rel_{Mtb} synthesis activity is not activated by deletion of the C-terminal, which may indicate a difference in Rel_{Mtb} and Relseq regulation.

[00263] The results presented herein demonstrate that the transferase and hydrolysis reactions of Rel_{Mtb} are catalyzed at distinct active sites. Rel_{Mtb} fragment 1-181 is only capable of (p)ppGpp hydrolysis, and protein fragment 87 to 394 is only capable of (p)ppGpp synthesis. While these activities function independently of one another in the absence of the RAC, the protein fragments share an overlapping region of amino acids 87 to 181 (Figure 1). This is also the case for E. coli SpoT and Relseq where an overlapping region is necessary for its (p)ppGpp synthesis and hydrolysis activities (10,12). The overlapping domain may contain residues involved in the mechanism of pyrophosphate transfer from either ATP or (p)ppGpp, which is also suggested by Mechold (12). The area may bind necessary divalent cations or may aid in the structural stability of the protein or assist in multimer formation. Alternatively the overlapping area may serve as a "cross talk" domain, between hydrolysis and synthesis domains, mediating a conformational change for the RAC-induced allosteric regulation of Rel_{Mtb} activity (6).

[00264] Much of this overlap occurs in a predicted region known as the HD domain, named for the conserved histidine and aspartate doublet in this region (Figure 5A) (13). This domain is found in a superfamily of enzymes with a predicted or known phosphohydrolase activity. Secondary-structure prediction shows that the conserved elements of the HD domain display a complex α/β structure (14). Based on catalytic properties of enzymes with this domain, the conserved H and D residues are predicted to be essential for coordinating divalent cations needed for activity. An HD domain in Rel_{Mtb} was identified using the PSI-BLAST program as amino acids 49-159 from the SMART database (14.4, 14.6). Since the entire HD domain is present in the 1-203 and 1-181 hydrolysis fragments, features of this domain may be necessary

for this reaction (Figure 5D,E). According to the SMART HD domain prediction, the 1-156 amino acid fragment has its HD domain disrupted (Figure 5F). Removal of the terminal amino acids from this domain may leave the fragment without necessary active site residues for (p)ppGpp hydrolysis.

[00265] Although much of the HD domain is present in the 87-394 (p)ppGpp synthesis only fragment, this Rel_{Mtb} fragment lacks the conserved histidine and aspartate doublet, and this may explain why it is not capable of (p)ppGpp hydrolysis, while the 1-394 fragment can both synthesize and hydrolyze (p)ppGpp (Figure 5G,C). Support for this hypothesis is the observation that E. coli RelA contains substitutions of the predicted histidine and aspartate doublet, which may account for the fact that it only has (p)ppGpp synthesis activity (10,13). In addition, mutation of the histidine residue within the HD signature of cGMP-specific phosphodiesterase produces a severe effect on its cGMP hydrolytic activity (15).

[00266] Gropp and colleagues found that the E. coli RelA mutation G251E caused a complete loss of RelA (p)ppGpp synthesis ability in culture, which agreed with the observation of the analogous spontaneous Bacillus subtilis mutant that exhibited a relaxed phenotype (16, 17). Gropp further found that a RelA H354Y mutation caused a partial loss of synthetic activity (16). RelA-G251E is impaired in its ability to bind both ATP and GTP, while RelA-H354Y is impaired in its ability to bind ATP but can still bind GTP, and it was suggested that both residues may be needed in the active site (16). RelA G251 and H354 are conserved in wild type Rel_{Mtb} and correspond to Rel_{Mtb} G241 and H344, respectively (Figure 5). Both residues are present in the Rel_{Mtb} 87-394 fragment and this may explain why it still has (p)ppGpp synthesis activity while the 1-203 and 1-181 fragments do not (Figure 5D,E,G).

[00267] Wild type Rel_{Mtb} synthesis activity is activated 2-fold from basal levels by uncharged tRNA alone and 80-fold by a complex of uncharged tRNA, ribosomes, and cognate mRNA (RAC) (6). The inability of the 1-394 and 1-450 Rel_{Mtb} synthesis fragments to be activated above basal levels by the RAC indicates that the C-terminal 288 amino acids contain an area of Rel_{Mtb} regulation (Figure 1). This agrees with work by Metzger and Schreiber, who found that the N-terminal domain (1-455) of RelA contains the catalytic domain, and the C-terminal domain (455-744) of RelA, which is devoid of synthetic activity, is involved in regulating activity (18,19). Yang has also reported that RelA amino acid region 455-C terminus represents the ribosome binding domain of the protein (20). This ribosomal binding area identified by Yang corresponds to the Rel_{Mtb} region 449-C terminus (Figure 5A).

Two conserved domains lie within the deleted C-terminal region of RelMtb:

[00268] The TGS and ACT domains. The TGS domain from 396-459 in Rel_{Mtb}, named for Threonyl-tRNA synthetase, GTPase, and SpoT is a small domain that consists of about 50 amino acid residues and is predicted to possess a predominantly beta-sheet structure (Figure 5A) (21). The function of the TGS domain is unknown, but Wolf suggests a nucleotide binding regulatory role because of its presence in two types of regulatory proteins (the GTPases and guanosine polyphosphate phosphohydrolases/synthetases) (21). The ACT domain, amino acids 661-734 in Rel_{Mtb}, is linked to a wide range of metabolic enzymes that are regulated by amino acid concentration (Figure 5A). Several enzymes are classic examples of allosteric regulation by the end products of the respective pathways (22). Pairs of ACT domains bind specifically to a particular amino acid leading to regulation of the linked enzyme. According to Aravind, the ACT domain may be a conserved, evolutionary mobile module that was independently fused to a variety of enzymes, which made these enzymes susceptible to regulation by amino acids (22). Thus Rel_{Mtb} activity may in part be regulated by ACT domain recognition of the amino acid on the terminal adenosine of charged tRNA or of specific amino acids free in the cell.

[00269] A variety of uncharged tRNAs bind Rel_{Mtb}, and the free 3'-OH of the ribosome ring of the terminal adenosine on the 3' end of uncharged tRNA is necessary for Rel_{Mtb} synthesis activation (6). Therefore it seems likely that a distinct domain of Rel_{Mtb} would recognize uncharged tRNA. Like the eukaryotic GCN2 protein, a stress responsive protein activated by uncharged tRNA, Rel_{Mtb} contains an amino acid domain with homology to histidyl-tRNA synthetase that enables it to bind a variety of uncharged tRNAs (23,26). Although Rel_{Mtb} amino acid sequence 400-738 has only 19% identity and 16% similarity to the entire sequence of Mtb histidyl-tRNA synthetase (HisS), important conserved residues, including residues in the m1, m2, and m3 signature motifs of class II aminoacyl-tRNA synthetases, are conserved in Rel_{Mtb} (Figure 5A). HisS is separated by only two genes (1,738 base pairs) from Rel_{Mtb}, suggesting gene duplication of HisS as a likely scenario for Rel_{Mtb}'s homology. The m1 motif of class II aminoacyl-tRNA synthetases has a role in the ability of the aminoacyl-tRNA synthetases to multimerize, and may have a similar role in Rel_{Mtb} (27). Interestingly, the TGS domain overlaps with the m1 motif of class II tRNA synthetases in Threonyl-tRNA synthetase and Rel_{Mtb} (Figure 5). A specific function for the TGS domain has not been identified, therefore we suggest that the TGS domain may have a role in multimerization. The m2 motif, which has been shown to interact with the acceptor stem of tRNA from crystal structure analysis of aspartyl-tRNA synthetase complexed with tRNAAsp (27), contains an invariant arginine that is conserved in Rel_{Mtb}, E. coli RelA, Mtb HisS, and all GCN2 proteins from yeast to humans. Mutation of this

conserved arginine (R1120L) in Saccharomyces cerevisiae GCN2 abolishes uncharged tRNA binding, suggesting its importance for conservation through evolution (25). It is possible that only a few amino acids in histidyl-tRNA synthetase are homologous to the C-terminal (400-738) area of Rel_{Mtb} because only essential amino acids capable of recognizing general uncharged tRNA structure may have been conserved in Rel_{Mtb}. In this way Rel_{Mtb} is non-discriminatory, binding a variety of uncharged tRNAs.

[00270] Rel_{Mtb} protein fragments 1-394 and 87-394 have a kcat/Km for basal level (p)ppGpp synthesis that is slightly lower than that for wild type Rel_{Mtb} (Table 4). In contrast, Relseq N-terminal fragment NH 1–385 had a (p)ppGpp synthesis specific activity that was activated 12-fold relative to wild type Relseq (12). There is a striking decrease in kcat/Km for basal (p)ppGpp hydrolysis by the Rel_{Mtb} truncated proteins versus that by the wild type Rel_{Mtb} (Table 5). Relseq also had decreased hydrolysis specific activity when the C-terminus of the protein was removed (12). Because Km values for (p)ppGpp hydrolysis and synthesis are similar for the Rel_{Mtb} fragment proteins and wild type Rel_{Mtb}, it is likely that substrate binding sites have not been disrupted (Table 4,5). The dramatic decrease in the kcat for (p)ppGpp hydrolysis and slight decrease in the kcat for (p)ppGpp synthesis by the protein fragments compared to that by wild type Rel_{Mtb} suggests that there are other regulatory domains in the C-terminal region of Rel_{Mtb}.

[00271] E. coli RelA activity is regulated by oligomerization mediated by the C-terminal domain of RelA (16). Amino acid regions 455-538 and 550-682 in RelA, corresponding to 449-535 and 547-676 in Rel_{Mtb} respectively, have been identified to take part in dimerization (Figure 5A) (20). Gropp's results from mutation analysis further suggest that amino acids C612, D637, and C638, which are part of a highly conserved 27 amino acid sequence in RelA homologues, are important for RelA dimerization (16). These three amino acids are conserved at positions C607, D632, C633 in Rel_{Mtb}, and Rel_{Mtb} has 67% homology to the conserved 27 amino acid sequence of RelA (Figure 5A).

[00272] The results from gel filtration assays indicate that wild type Rel_{Mtb} is capable of forming trimers (Figure 4A). Incubation with DTT did not disrupt the trimer, suggesting that disulfide bonds are not responsible for oligomerization. The 1-394 fragment is capable of forming monomers and trimers, indicating that there are still amino acids in the C-terminal region of this fragment allowing trimer formation or the N-terminal half of Rel_{Mtb} alone is capable of forming multimers (Figure 4B). Gropp found RelA C-terminal to C-terminal interactions and he also found weak interactions of full-length RelA with the N-terminus of RelA

(16). We found only a single peak corresponding to the monomer state for the 1-181 (Figure 4C) and 1-203 fragments (Figure 4D), therefore we suggest that the amino acid area of 203-738 is the major regulator of Rel_{Mtb} multimer formation. Because we could not isolate monomers of wild type Rel_{Mtb}, it appears that wild type Rel_{Mtb} multimer equilibrium highly favors the trimer state over the monomer state.

[00273] Analysis of isolated 1-394 trimer and isolated 1-394 monomer activity indicates that one monomer has the same activity as one trimer for (p)ppGpp synthesis. One of the three subunits of the trimer may be functional or three subunits of the trimer may form a new catalytic site. Since the KGTP is unchanged for synthesis, we suggest that two components of the 1-394 trimer are inactive for (p)ppGpp synthesis. The 1-394 monomer and trimer carry out decreased basal levels of (p)ppGpp synthesis and hydrolysis compared to wild type Rel_{Mtb}. While capable of trimerization, the 1-394 fragment may not associate in the same way as wild type Rel_{Mtb}. The C-terminus (395-738) of Rel_{Mtb} may contain binding domains that are necessary to keep the trimer in a proper conformation for basal catalytic function. This is similar to the case for yeast class II aspartyl-tRNA synthetase, where an invariant proline in the motif I multimerization domain is necessary as a structural determinant for proper subunit interface association and resulting catalytic activity (33).

[00274] C-terminal overexpression of E. coli RelA inhibits (p)ppGpp accumulation in relA+ cells starved for amino acids (16), suggesting a multimeric state is an inhibitory state for (p)ppGpp synthesis. Rel_{Mtb} 1-394 monomers only had a three-fold activity increase over Rel_{Mtb} 1-394 trimers and neither the 1-394 monomers nor trimers had basal (p)ppGpp hydrolysis or synthesis activity. Multimerization alone does not appear to be responsible for Rel_{Mtb} activity regulation. We propose the Rel_{Mtb} C-terminal end may loop back to its N-terminus and interact with a site important for regulating (p)ppGpp synthesis and hydrolysis, which is similar to the case for GCN2. An interesting possibility would be that the Rel_{Mtb} C-terminus binds to the overlapping hydrolysis and synthesis domain (amino acids 87-181), utilizing this "cross talk" area for regulation.

[00275] The RAC is important for regulating the catalytic activities of Rel_{Mtb} (6), so it is likely that components of the RAC affect domain interactions that result in (p)ppGpp synthesis and hydrolysis alterations. Alone, uncharged tRNA can increase the kcat/Km for synthesis 2-fold and decrease the kcat/Km for hydrolysis 5-fold (6). The N-terminus of Ribosomal Protein L11, which is involved in regulating the activities of the GTPases EF-Tu and EF-G, is involved in regulation of the RelA protein in E. coli (32). Mutations that map to ribosomal protein L11

(relC mutations) confer a relaxed phenotype (34,35). RAC components are likely to affect multimerization and C-terminal-"cross talk" interactions.

[00276] Wild type Rel_{Mtb} was found to exclusively form trimers, so we suggest that Rel_{Mtb} maintains a trimer state during normal growth condition. A multimer state may be needed for Rel_{Mtb} to bind uncharged tRNA, as is the case for GCN2 (31), so Rel_{Mtb} trimerization may be an environmentally responsive state. In our proposed model, the trimerized Rel_{Mtb} will be bound to the ribosome and carry out basal level (p)ppGpp synthesis and hydrolysis. Upon amino acid starvation, uncharged tRNA enters the A site of the ribosome and binds to the m2 motif of Rel_{Mtb}. Binding of uncharged tRNA and involvement of the L11 ribosomal protein (32) would cause the 80 fold increase in kcat/Km for synthesis (6) and the 6 fold decrease in kcat/Km for hydrolysis (6) by alterating C-terminal binding to the "cross talk" area and dissociating the trimer. Crystal structure and NMR analysis and investigation of uncharged tRNA binding and ribosomal interaction with Rel_{Mtb} will elucidate this mechanism further.

[00277] Response to the problem of amino acid starvation seems to have been solved in both prokaryotes and eukaryotes with similar structural components. Upon activation by uncharged tRNA, the prokaryotic Rel protein and the eukaryotic GCN2 protein increase amino acid biosynthetic enzymes (Figure 6). Both proteins are similarly organized; their N-terminal regions contain a catalytic domain while their C-terminal regions are involved in regulation involving uncharged tRNA binding, ribosomal binding, and multimerization. While differences exist, the "general control" response of GCN2 (28,29), first identified in Saccharomyces cerevisiae, may be an evolutionarily modified version of the Rel "stringent response" (30) of bacteria. Future work identifying structural similarities and regulation properties between Rel, GCN2, and their pathways may uncover the evolution of a single survival response to amino acid depravation that has been conserved from bacteria to mammals.

Example 7 Effect of Rel_{Mtb} on Transcription.

[00278] The overall number of genes regulated by Rel_{Mtb} is comparable to that observed for RelA of other bacteria. This suggests that other, non-(p)ppGpp-dependent parallel regulons are engaged following starvation.

[00279] A strain of Mtb in which the gene encoding Rel_{Mtb} has been allelically disrupted does not make (p)ppGpp during starvation conditions and is dramatically reduced in its ability to survive in long-term cultures.

[00280] We used a murine infection model to show that the Δ Rel_{Mtb} strain can establish an infection in the lungs and spleens of mice but significantly attenuated in its ability to persist

relative to the WT strain. Correspondingly, in some embodiments the ΔRel_{Mtb} strain can provide immune protection against infection with wild type Mtb. Comparison of the transcriptomes of the WT and ΔRel_{Mtb} strains cultured under stringent conditions revealed alterations in transcripts encoding gene products involved in bacterial survival including protein translation, DNA replication, nucleotide metabolism, and lipid metabolism. The results presented here support our model that eliminating the major bacterial mechanism for starvation adaptation impairs persistence in the host. Therefore, ΔRel_{Mtb} allows the study of a wide range of cellular and biochemical processes in a genetically well defined host strain and may form the basis of a live attenuated vaccine candidate. Deleting the entire Rel_{Mtb} gene, or regulatory and catalytic domains of this gene described herein, could produce an attenuated strain without the ability to attain dormancy.

[00281] It is appreciated further that the inventive concepts described herein may be extended to produce various vaccines wherein the Rel Mtb gene is removed in combination with one or more additional genes of mycobacteria as described below. Furthermore, various drugs may be designed to inhibit the creation of the Rel Mtb gene alone or in combination with one or more additional gene. Lastly, it is appreciated that vaccines and drugs may be produced wherein the vaccine removes one ore more of the below described list of genes and wherein the drug inhibits the formation of the gene within the mycobacteria.

[00282] Specifically, genome-driven screening methods for genes essential for virulence of MTB in mice have been the subject of intense study since the appearance of the complete genome sequence of MTB H37Rv28. Signature-tagged mutagenesis (STM) of MTB has been applied by two different laboratories to collectively discover 19 attenuated MTB mutants, yet neither group reported Rel_{Mtb} as a gene required for productive murine infection(29,30). The failure to identify Rel_{Mtb} may have been due to poor saturation of the technique.

[00283] However, a newly developed alternative approach integrates saturating mutagenesis and microarrays to more exhaustively identify genes essential for growth on minimal media(31,32) confirmed that Rel_{Mtb} is essential for bacterial survival during infection (Rubin, E., Sassetti, C., personal communication).

[00284] Recently, a whole genome microarray analysis of H37Rv under severe starvation conditions- transfer of cultures in phosphate buffered saline- was reported(6). Although in general, concordance between this data and our current data set was high, this analysis revealed many more genes regulated upon entry into stationary phase than we report to be controlled by Rel_{Mtb}, perhaps reflecting the different growth conditions that were used to induce the Rel

dependent genes. The overall number of genes regulated by Rel_{Mtb} is comparable to that observed for RelA of other bacteria. This suggests that other, non-(p)ppGpp-dependent parallel regulons are engaged following starvation.

[00285] Here we demonstrate that Rel_{Mtb} regulates several genes that have been implicated in, or can plausibly be linked directly to the virulence of MTB. Resuscitation promoting factor (Rpf) is a protein widely distributed among Gram-positive bacteria that has been shown to facilitate the recovery of dormant Micrococcus luteus, Rhodococcus rhodochrous, and MTB cells(33). MTB has five homologs of Rpf and one of these, rpfC (Rv1884c), is shown here to be strongly upregulated in a Rel_{Mtb}-dependent manner. Rpf-containing supernatants of spent cultures of MTB facilitate resuscitation of dormant MTB cells suggesting that this gene product exerts an effect early in the growth cycle. Nonetheless, dormant cells may not have protein synthetic capacity and cells may need to be "primed" with an Rpf precursor which could be processed or released independently of new protein synthesis for maximal viability upon reintroduction of nutrients. Although only picomolar concentrations of this molecule are required for resuscitation activity, the active form of the molecule has not been identified.

[00286] MTB contains several protein families composed largely of repeated amino acid motifs that have been proposed to play a role in antigenic variation among different clinical isolates28,34. Two members of the PPE family, three members of the PE family and one of the PE-PGRS family were identified here as being differentially regulated by Rel_{Mtb}. In all but one case, (PE20, Rv1806) Rel_{Mtb} activity serves to upregulate expression of these proteins upon starvation. Although the Function: S of these genes remain a matter of speculation, it is possible that they either provide antigenic diversity or inhibit antigen processing. Evidence that PE-PGRS proteins may play a role in virulence was shown by the deletion of two members of the PE-PGRS family from M. marinum that impaired the ability of this pathogen to persist in granulomas in frogs(16)

[00287] Transformation of a library of DNA from MTB into the saprophyte Mycobacterium smegmatis followed by enrichment by passage through macrophages demonstrated that the presence of Rv2416c (eis) significantly enhanced survival(25). Although the molecular function of this gene and its protein product are not known both protein and gene appear to be strongly downregulated by Rel_{Mtb}.

[00288] Several other genes were upregulated in a Rel_{Mtb}-dependent manner. These include Rv3588c (carbonic anhydrase) and Rv2287 (yjcE, probable Na+/H+ exchanger) both of which may confer relative resistance to phagosome acidification, and the hupB gene (Rv2986c) which

is homologous to E. coli HU protein that Function: S to stabilize DNA under extreme environmental conditions and to play a crucial role in stationary phase survival35. Proteins of interest that are downregulated by Rel_{Mtb} include; hbhA (Rv0475) that mediates MTB binding to epithelial cells may contribute to extrapulmonary dissemination36, fbpA (Rv3804c) that encodes for a member of the Ag85 complex involved in mycolic acid transfer to the growing mycobacterial cell wall, and mpt83 (Rv2873) that encodes a cell-surface lipoprotein. The whm (whiB homologs in mycobacteria) family of genes encode sigma factors homologous to a gene known to be involved in sporulation of Streptomyces coelicolor and consists of six members in MTB. Because of the tantalizing similarities between sporulation and latency, the first whm gene discovered (whiB3) was studied in detail but was found not to be induced in stationary phase, nor upon oxygen deprivation.

[00289] Rather than a direct effect on expression of a known virulence factor there is some evidence that the loss of viability in $H37Rv\Delta Rel_{Mtb}$ reflects better immune recognition of this strain due to dysregulation of immunomodulatory bacterial antigens. This is an attractive hypothesis, in part, because of the decreased amount of inflammation observed upon histopathologic examination of mouse lungs in this study. Among the more potent bacterial antigens found to be upregulated by Rel_{Mtb} was esat6 (Rv3875). Deletion of this gene from Mycobacterium bovis produces a strain that is attenuated in guinea pigs37. In addition to its biochemical role in cell wall formation, Antigen 85A has been shown to drive naïve human T cells to differentiate towards a Th1 phenotype in *in vitro* models(38). When Antigen 85A is injected into C57BL/6 mice an elevated IFN-□ response is generated that is protective against intravenous challenge with wild type MTB(39). Thus the relative overexpression of Antigen 85A in the mutant strain which fails to downregulate it may provide for enhanced clearance due to increased host recognition of this protein. The 10-kD chaperone protein of MTB, GroES, was also found to be down-regulated in a Rel_{Mtb}-dependent manner and is highly immunostimulatory (40). Likewise Mpt83 (Rv2873), Mpt70 (Rv2875), GroEL2 (Rv0440), and eis (Rv2416c) are downregulated specifically by Rel_{Mtb} and are known to be highly immunogenic(41).

[00290] The results of these experiments show that the MTB stringent response is an element that is useful for bacterial survival under persistent infection conditions in mice. Rel_{Mtb} is a global regulator of gene expression that activates the protective response when the organism is under physiologic and immunologic stress. Genes regulated in a Rel_{Mtb}-dependent manner have been identified, and this regulon provides a critical list of potential chemotherapeutic targets ideally suited for developing novel treatments for latent infections or for shortening the course of

antituberculosis chemotherapy(6). Rel_{Mtb}-regulated genes also provide a rich source of potentially specific diagnostic markers that may be predictive of successful tuberculosis chemotherapy.

RelMtb downregulates the mycobacterial translational apparatus in response to starvation

[00291] To begin to identify the specific gene products that determine the *in vivo* phenotype of the H37RvΔRel_{Mtb} strain, we compared the transcriptional response of this strain with that of the WT and complemented strains during nutrient starvation. Whole-genome microarray analysis was performed on both strains starved for nutrients, and genes that were differentially regulated were identified (see Methods). WT H37Rv downregulates the cellular translational apparatus, consistent with Rel-mediated effects in other bacteria. There are 59 ribosomal proteins encoded within the genomic DNA sequence of MTB. The oligonucleotide arrays utilized in this experiment included probes for 50 of the encoding genes, but only 34 produced fluorescence intensities detectable in normally growing cells. Of these, 15 genes (44% of those interrogated by this method) appeared in the filtered list of differentially regulated genes controlled by Rel_{Mtb}. Other elements of the translational apparatus, such as EF-Tu (Rv0685) and initiation factor IF-1 (Rv3462c) also appeared to be downregulated by engagement of Rel_{Mtb}. The operonic structure of many of the ribosomal proteins suggests that genes within a cluster should be coordinately-regulated. Indeed, many examples of ribosomal genes within a cluster with down-regulated members were repressed but failed to survive the rigorous filters applied to the data sets (see Supplementary Information for the complete data sets). Stable RNAs are not represented in the microarray data because oligos for rRNAs and tRNAs were not included in the oligo set used.

RelMtb regulates the expression of known secreted antigens and virulence factors

[00292] In all, 70 genes encoding proteins not directly involved in protein synthesis appear to be Rel_{Mtb}-regulated in the microarray experiment (Table 6). Rel_{Mtb}-regulated genes include the following: hupB (Rv2986c) that encodes a DNA binding protein, mas (Rv2940c) that encodes mycocerosic acid synthase, genes for chaperonin or heat shock proteins such as groEL2 (Rv0440), Rv0251c, and groES (Rv3418c), genes for cell wall biosynthetic proteins such as RmlB and antigen 85A, and a gene homologous to the sporulation sigma factor gene whiB6 (Rv3862c). Rel_{Mtb}-regulated genes also included a number of known or suspected virulence factors including eis (Rv2416c), Rv3588c (carbonic anhydrase), Rv2287 (yjcE, probable Na+/H+ exchanger), rpfC (Rv1884c), genes encoding six members of the PE/PE-PGRS protein family (Rv3508, Rv1806, Rv2431c, Rv1195), and hbhA (Rv0475).

[00293] Five known secreted antigens were found to be regulated by Rel_{Mtb} including ESAT-6 (Rv3875), Ag85A (Rv3804c), Rv0475, Mpt83 (Rv2873), and Mpt70 (Rv2875). ESAT-6 was upregulated in the WT strain induced by Rel_{Mtb} for the stringent response but not in the ΔRel_{Mtb} strain. The remaining secreted antigens were downregulated in the wild type strain compared to the ΔRel_{Mtb} strain.

[00294] Independent confirmation of the microarray data was performed by Western Blot analysis of lysates derived from H37Rv and H37RvΔRel_{Mtb} grown to stationary phase in rich media. Primary antibodies specific for the Eis protein (Rv2416c)25 indicated accumulation of this protein in the ΔRel_{Mtb} strain but not in the WT strain, confirming the downregulation of eis seen in Tables 6.

[00295] **Tables 6 and 7:** Microarray comparisons of WT and Δrel MTB strains induced for the stringent response. Genes are listed based upon their Rv synonyms of the annotated MTB genome. Names for genes with known or suspected function can be found by entering the Rv numbers in the TubercuList web server (http://genolist.pasteur.fr/TubercuList/). Tuberculist has divided all putative MTB gene products into 9 functional categories : 0, virulence, detoxification, adaptation; 1, lipid metabolism; 2, information pathways; 3, cell wall and cell processes; 4, stable RNAs; 5, insertion sequences and phages; 6, PE and PPE family members; 7, intermediary metabolism and respiration; 8, unknown; 9, regulatory proteins; and conserved hypototheticals. Genes with differential expression between the WT and Δrel MTB strains are grouped into these 9 categories in the table. Differences in expression levels are demarked as +, – = 1.85-2.0 fold differences; ++, – – = 2.01-3.0 fold differences; and +++, – – – = ≥ 3.01 fold differences

TubercuList functional category	Rv number	Relative expression in WT strain	TubercuList functional category	Rv number	Relative expression in WT strain
0	0440	– – –	7	3138	+ +
	3418c	– – –		3340	– – –
	0251c	– –		3029c	– –
1	2940c	+		2725c	– –
	3804c	– –		3464	– –
2	0706	+ +		2584c	– –
	0683	– –		3588c	+ +
	3461c	– – –		3762c	+ +
	3459c	– –		2535c	+ +
	3460c	– –	8	3889c	+ +

	2441c	– –
	3443c	– – –
	0700	– – –
	3462c	– –
	0682	– –
	0701	– –
	2890c	– –
	0714	– –
	3924c	– –
	0685	– –
	0716	– –
	2442c	– –
	0641	–
	2986c	+ +
	0054	– –
3	3894c	+ +
	0847	– –
	2287	+ +
	0475	– –
	3875	+ +
	2873	– – –
	2875	– –
	1884c	+ + +
4	no	genes
5	2791c	+ + +
	2657c	– –
6	1196	+ +
	3478	+ +
	2431c	+ +
	1195	+ + +
	3508	+
	1806	– –
	2827c	– –
	0333	– –
	1955	– –
	0333	– –
9	1404	– –
	2034	– – –
	2989	– – –
	3863c	– –
	3862c	– –
10	3463	– –
	0288	+ +
	1810	+ +
	3876	+ +
	3182	– –
	1284	– –
	3385c	– –
	2416c	– – –
	1103c	– – –
	2641	– –
	2239c	– –
	3615c	– –
	0581	– –
	0190	– –
	2295	– –
	0581	– –
	3182	– – –
	2147c	– –
	2632c	– –

Process	Rv number	gene name (description)	expression in WT strain	Notes
Transcription	0706	*rplV*	++	1
	0683	*rpsG* (30S ribosomal protein [S7])	—	2
	3461c	*rpmJ* (50S ribosomal protein [L36])	——	3
	3459c	*rpsK* (30S ribosomal protein [S11])	—	4
	3460c	*rpsM* (30S ribosomal protein [S13])	—	5
	1404	(ArsR repressor protein)	—	6
	2441c	*rpmA* (50S ribosomal protein [L27])	—	7
	2034	*arsR* family (transcriptional regulator)	——	8
	3443c	*rpmL* (50S ribosomal protein [L13])	——	9
	2989	(probable transcriptional regulator)	——	10
	0700	*rpsJ* (30S ribosomal protein [S10])	——	11

	3462c	*infA* (initiation factor IF-1)	—	12
	0682	*rpsL* (30S ribosomal protein [S12])	—	13
	0701	*rplC* (50S ribosomal protein [L3])	—	14
	2890c	*rpsB* (30S ribosomal protein [S2])	—	15
	0714	*rplN* (50S ribosomal protein [L14])	—	16
	3924c	*rpmH* (50S ribosomal protein [L34])	—	17
	0685	*tuf* (elongation factor EF-Tu)	—	18
	0716	*rplE* (50S ribosomal protein [L5])	—	19
	2442c	*rplU* (50S ribosomal protein [L21])	—	20
	0641	*rplA* (50S ribosomal protein [L1])	-	21
Cellular processes (other than Transcription)	2986c	*hup* (DNA binding protein)	++	22
	2791c	(transposase)	+++	23
	3138	*pflA* (pyruvate lyase activating protein)	++	24
	3894c	(transmembrane ATP/GTP binding protein)	++	25
	2940c	*mas* (mycocerosic acid synthase)	+	26
	0440	*groEL2* (60 kD chaperonin 2)	——	27
	3340	*metC* (cstathionine b-lyase)	——	28
	0251c	*hsp* (heat shock protein)	—	30
	0847	*lpqS* (lipoprotein)	—	31
	0054	*ssb* (single stgrand binding protein)	—	32
	3029c	*fixA* (electron transfer flavoprotein b sub.)	—	33
	2725c	*hflx* (GTP binding protein)	—	34
	3463		—	35
	3464	*rmlB* (dTDP-glucose 4,6-dehydratase)	—	36
	2584c	*apt* (adenine phosphoribodyltranserase)	—	37
	3588c	(carbonic anhydrase)	++	38
	2287	*yjcE* (probable Na^+/H^+ exchanger)	++	39
	3762c		++	40
	2535c	*pepQ* (cytoplasmic peptidase)	++	41
Secreted proteins	0475	*hbhA*	—	42
	3875	*esat-6*	++	43
	3804c	*fbpA* (Ag85A [mycolyltransferase])	—	44
	2873	*mpt83* (surface lipoprotein)	——	45
	2875	*mpt70* (major secreted immunogenic protein)	—	46
	3418c	*groES* (10 kD chaperone)	——	47
PE and PPE Family members	1196	PPE18	++	48
	3478	PPE60	++	49
	2431c	PE25	++	50
	1195	PE13	+++	51
	3508	PE PGRS54	+	52
	1806	PE20	—	53
Unknown	0288		++	54
	1810		++	55

	3889c		++	56
	3876		++	57
	3182		—	58
	2827c		—	59
	0333		—	60
	1284		—	61
	3385c		—	62
	2416c	*eis*	——	63
	1103c		——	64
	2657c	Probable phiRv2 prophage protein	—	65
	2641	*cadI* (cadmium inducible protein)	—	66
	2239c		—	67
	3615c		—	68
	1955		—	69
	1884c	*rpfC* (bacterial cytokine)	+++	70
	0581		—	71
	0190		—	72
	2295		—	73
	0333		—	74
	0581		—	75
	3182		——	76
	2147c		—	77
	3862c		—	78
	2632c		—	79

[00296] These genes are better described by the following, wherein the gene's function is described by:

0 virulence, detoxification, adaptation

1 lipid metabolism

2 information pathways

3 cell wall and cell processes

4 stable RNAs

5 insertion seqs and phages

6 PE/PPE

7 intermediary metabolism and respiration

8 unknown

9 regulatory proteins

10 conserved hypotheticals

[00297] **Table 7:** Description of Genes whose Expression is Modified in Rel Mutant

1. Function this protein binds specifically to 23S rRNA; its binding is stimulated by other ribosomal proteins, e.g., L4, L17, AND L20. It is important during the early stages of 50s reconstitution. Product: Probable 50S RIBOSOMAL PROTEIN L22 RPLV Comments: Rv0706, (MTCY210.25), length: 197 aa. Probable rplV, 50S ribosomal protein L22 (information pathways 2)

2. PROTEIN S7 BINDS SPECIFICALLY TO PART OF THE 3' END OF 16S RIBOSOMAL RNA. Product: Probable 30S RIBOSOMAL PROTEIN S7 RPSG Comments: Rv0683 (2)

3. Function: Involved in translation mechanism. Product: Probable 50S RIBOSOMAL PROTEIN L36 RPMJ Comments: Rv3461c, (MTCY13E12.14c), length: 37 aa. Probable rpmJ, 50S ribosomal protein L36, (2)

4. Function: S11 plays an essential role for the selection of the correct tRNA in protein biosynthesis. It is located on the large lobe of the small subunit. Product: Probable 30S RIBOSOMAL PROTEIN S11 RPSK Comments: Rv3459c (2)

5. Function: Involved in the binding of fMET-tRNA and, hence, in the initiation of translation. Product: Probable 30S RIBOSOMAL PROTEIN S13 RPSM Comments: Rv3460c, (MTCY13E12.13c), length: 124 aa. Probable rpsM, 30S ribosomal protein S13 (2)

6. Involved in transcriptional mechanism Product: Probable TRANSCRIPTIONAL REGULATORY PROTEIN (9)

7. Involved in translation mechanisms. Product: Probable 50S RIBOSOMAL PROTEIN L27 RPMA (2)

8. Function: unknown Product: Probable ArsR repressor protein Comments: Rv2034, (MTV018.21), length: 107 aa. Probable repressor protein similar to several belonging to the ARSR FAMILY *e.g.* Q53040 (112 aa) (regulatory proteins 9)

9. Function: Involved in translation mechanism. this protein is one of the early assembly proteins of the 50s ribosomal subunit. Product: Probable 50S RIBOSOMAL PROTEIN L13 (2)

10. Function: Involved in transcriptional mechanism. product: probable transcriptional regulatory protein Evidence: experimental Proteomics: Identified by proteomics at the Max-Planck-Institut fuer Infektionsbiologie (Germany) (see citation below). Comments: Rv2989, (MTV012.03), length: 233 aa. Probable transcriptional regulator (ala-rich protein), highly similar to O86533|SC1C2.33c PUTATIVE TRANSCRIPTIONAL REGULATOR from Streptomyces coelicolor (238 aa), FASTA scores: opt: 711, E(): 2.3e-38, (53.05% identity in 230 aa overlap); and similar to others *e.g.* Q9KND6 PUTATIVE TRANSCRIPTIONAL REGULATOR from Vibrio cholerae (244 aa), FASTA scores: opt: 232, E(): 1.2e-07, (29.75% identity in 232 aa

overlap); Q9R9U0|SRPS EFFLUX PUMP REGULATOR from Pseudomonas putida (259 aa), FASTA scores: opt: 224, E(): 4.1e-07, (28.35% identity in 247 aa overlap); etc. Also similar to proteins from Mycobacterium tuberculosis *e.g.* O06806|Rv1773c|MTCY28.39 HYPOTHETICAL 26.6 KDA PROTEIN (248 aa), FASTA scores: opt: 239, E(): 4.4e-08, (29.85% identity in 231 aa overlap); P71977|RV1719|MTCY04C12.04 HYPOTHETICAL 27.9 KDA PROTEIN (259 aa), FASTA scores: opt: 215, E(): 1.6e-06, (31.85% identity in 223 aa overlap); etc. Equivalent to AAK47396 from Mycobacterium tuberculosis strain CDC1551 (267 aa) but shorter 34 aa. Contains possible helix-turn-helix motif at aa 25-46 (Score 1005, +2.61 SD). TBparse score is 0.894. Jungblut PR, Schaible UE, Mollenkopf HJ, Zimny-Arndt U, Raupach B, Mattow J, Halada P, Lamer S, Hagens K, Kaufmann SH. Comparative proteome analysis of Mycobacterium tuberculosis and Mycobacterium bovis BCG strains: towards functional genomics of microbial pathogens Mol Microbiol (1999) 33:1103-17 (9)

11. This protein is involved in the binding of tRNA to the ribosomes Product: Probable 30S RIBOSOMAL PROTEIN S10 RPSJ Comments: Rv0700, (MTCY210.19), length: 101 aa. Probable rpsJ, 30S ribosomal protein S10, equivalent to RS10_MYCLE P307653 30S ribosomal protein S10 from Mycobacterium leprae (101 aa), FASTA scores: opt: 645, E(): 0, (97.0% identity in 101 aa overlap). Also highly similar to others *e.g.* CAB82069.1|AL161803 30S ribosomal protein S10 from Streptomyces coelicolor (102 aa); etc. Contains PS00361 Ribosomal protein S10 signature. Belongs to the s10p family of ribosomal proteins. (2)

12. No specific function has so far been attributed to this initiation factor; however, it seems to stimulate more or less all the activities of the other two initiation factors, IF-2 AND IF-3. Product: Probable TRANSLATION INITIATION FACTOR IF-1 INFA Comments: Rv3462c, (MTCY13E12.15c), length: 73 aa. Probable infA, initiation factor IF-1, equivalent to P45957|ML1962|INFA TRANSLATION INITIATION FACTOR IF-1 from Mycobacterium bovis (72 aa) and Mycobacterium leprae (72 aa), FASTA scores: opt: 472, E(): 6.6e-28, (100.0% identity in 72 aa overlap). Also highly similar to others *e.g.* O54209|IF1_STRCO|INFA|SC6G4.03 from Streptomyces coelicolor (73 aa), FASTA scores: opt: 424, E(): 2e-24, (84.95% identity in 73 aa overlap); O50630|IF1_BACHD|INFA|BH0158 from Bacillus halodurans (71 aa), FASTA scores: opt: 388, E(): 8.1e-22, (77.8% identity in 72 aa overlap); Q9XD14|IF1_LEPIN|INFA from Leptospira interrogans (71 aa), FASTA scores: opt: 376, E(): 6e-21, (80.0% identity in 70 aa overlap); etc. Contains 1 S1 motif domain. Belongs to the if-1 family. (2)

13. Function: protein s12 is involved in the translation initiation step. Product: Probable 30S RIBOSOMAL PROTEIN S12 (2)

14. This protein binds directly to 23s ribosomal rna and may participate in the formation of the peptidyltransferase center of the ribosome. Product: Probable 50S RIBOSOMAL PROTEIN L3 (2)

15. Involved in translation mechanism. Product: Probable 30S RIBOSOMAL PROTEIN S2 RPSB Evidence: experimental Proteomics: Identified by proteomics at the Statens Serum Institute (Denmark) (see citation below). Rosenkrands I, King A, Weldingh K, Moniatte M, Moertz E, Andersen P. Towards the proteome of Mycobacterium tuberculosis. Electrophoresis (2000) 21:3740-56 (2)

16. This protein binds directly to 23s ribosomal RNA. Product: Probable 50S RIBOSOMAL PROTEIN L14 RPLN Evidence: experimental Proteomics: Identified by proteomics (see citation below). Rosenkrands I, King A, Weldingh K, Moniatte M, Moertz E, Andersen P. Towards the proteome of Mycobacterium tuberculosis. Electrophoresis (2000) 21:3740-56 (2)

17. Involved in translation mechanism. This protein is one of the early assembly proteins of the 50s ribosomal subunit (by similarity). Product: 50S RIBOSOMAL PROTEIN L34 (2)

18. Function: this protein promotes the GTP-dependent binding of aminoacyl-trna to the a-site of ribosomes during protein biosynthesis. Product: Probable ELONGATION FACTOR TU TUF (EF-TU) Evidence: experimental Proteomics: The putative product of this CDS corresponds to spots 1_367, 1_362, 1_171 and 1_349 identified in culture supernatant by proteomics at the Max-Planck-Institut fuer Infektionsbiologie (Germany), and spot 0685 (or EF-tu) identified in cell wall by proteomics at the Statens Serum Institute (Denmark) (see citations below). Mollenkopf HJ, Jungblut PR, Raupach B, Mattow J, Lamer S, Zimny-Arndt U, Schaible UE, Kaufmann SHA dynamic two-dimensional polyacrylamide gel electrophoresis database: the mycobacterial proteome via Internet. Electrophoresis (1999) 20:2172-80 Jungblut PR, Schaible UE, Mollenkopf HJ, Zimny-Arndt U, Raupach B, Mattow J, Halada P, Lamer S, Hagens K, Kaufmann SH. Comparative proteome analysis of Mycobacterium tuberculosis and Mycobacterium bovis BCG strains: towards functional genomics of microbial pathogens Mol Microbiol (1999) 33:1103-17 (2)

19. This is one of 3 proteins that mediate the attachment of the 5s rna into the large ribosomal subunit. Product: Probable 50S RIBOSOMAL PROTEIN L5 (2)

20. Involved in translation mechanisms. Product: Probable 50S RIBOSOMAL PROTEIN L21 (2)

21. Function: this protein binds directly to 23s ribosomal rna and is located in the neighborhood of the site where elongation factor TU is bound to the ribosome. Product: Probable 50S RIBOSOMAL PROTEIN (2)

22. This protein belongs to the histone like family of prokaryotic DNA-binding proteins which are capable of wrapping DNA to stabilize it, and prevent its denaturation under extreme environmental conditions. Product: Probable DNA-BINDING PROTEIN HU HOMOLOG HUPB (HISTONE-LIKE PROTEIN) (HLP) (21-KDA LAMININ-2-BINDING PROTEIN) Probable hupB (alternate gene names: hup, hlp, lbp21), DNA-binding protein HU homolog (resembles fusion between HU and histone), equivalent to others from Mycobacteria *e.g.* (2)

23. Required for the transposition of the insertion element is1602. Product: Probable TRANSPOSASE Comments: Rv2791c, (MTV002.56c), length: 459 aa. Probable IS1602 transposase for IS1602 element, similar to many *e.g.* (5 – insertion sequences and phages)

24. Function: Involved in cellular metabolism [catalytic activity: S-adenosyl-L-methionine + dihydroflavodoxin + [formate acetyltransferase]-glycine = 5'-deoxyadenosine + methionine + flavodoxin + [formate acetyltransferase]-glycine-2-yl radical]. Product: Probable Pyruvate formate lyase activating protein pFLA (formate acetyltransferase activating enzyme) ([pyruvate formate-lyase] activating enzyme) (7 - intermediary metabolism and respiration)

25. Unknown Product: possible conserved membrane protein (Functional category 3 - cell wall and cell processes)

26. Catalyzes the elongation of n-fatty acyl-coa with methylamalonyl-coa (not malonyl-coa) as the elongating agent to form mycocerosyl lipids. Product: Probable MULTIFUNCTIONAL MYCOCEROSIC ACID SYNTHASE MEMBRANE-ASSOCIATED MAS Azad AK, Sirakova TD, Rogers LM, Kolattukudy PE. Targeted replacement of the mycocerosic acid synthase gene in Mycobacterium bovis BCG produces a mutant that lacks mycosides Proc Natl Acad Sci USA (1996) 93:4787-92 Mathur M, Kolattukudy PE Molecular cloning and sequencing of the gene for mycocerosic acid synthase, a novel fatty acid elongating multifunctional enzyme, from Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guerin. J Biol Chem (1992) 267:19388-95 (Functional category 1 lipid metabolism)

27. Prevents misfolding and promotes the refolding and proper assembly of unfolded polypeptides generated under stress conditions. Product60 KDA CHAPERONIN 2 GROEL2 (PROTEIN CPN60-2) (GROEL PROTEIN 2) (65 KDA ANTIGEN) (HEAT SHOCK PROTEIN 65) (CELL WALL PROTEIN A) (ANTIGEN A) Evidence: experimental Proteomics: The product of this CDS corresponds to spots 1_212, 1_216, 1_233, 1_224, 1_247, 1_249, 1_25,

1_456, 1_457, 1_458 and 1_463 identified in culture supernatant by proteomics at the Max-Planck-Institut fuer Infektionsbiologie (Germany), and spots GroEL identified in cell wall and cytosol by proteomics at the Statens Serum Institute (Denmark) (see second to fifth citations below). Rv0440, (MTV037.04), length: 539 aa. groEL2 (alternate gene names: groL2, groEL-2, hsp65, hsp60), 60 kDa chaperonin 2 (see first citation below). Purified 65 kDa antigen can elicit a strong delayed-type hypersensitivity reaction in experimental animals infected with *M. tuberculosis*. This protein is one of the major immunoreactive proteins of the mycobacteria. This antigen contains epitopes that are common to various species of mycobacteria. Shinnick TM. The 65-kilodalton antigen of Mycobacterium tuberculosis. J Bacteriol (1987) 169:1080-1088 (functional category 0 - virulence, detoxification, adaptation)

28. Function: transforms o-acetylhomoserine into l-methionine [catalytic activity: o-acetyl-l-homoserine + methanethiol = l-methionine + acetate]. product: probable o-acetylhomoserine sulfhydrylase metc (homocysteine synthase) (o-acetylhomoserine (thiol)-lyase) (oah sulfhydrylase) (o-acetyl-l-homoserine sulfhydrylase) (7)

29. Rv number doesn't exist.

30. Unknown. Product: Probable HEAT SHOCK PROTEIN HSP Comments: Rv0251c, (MTV034.17c), length: 159 aa. Probable hsp, heat shock protein belonging to HSP20 family. Also similar to P30223|14KD_MYCTU 14 KDA ANTIGEN (16 KDA ANTIGEN) 19K major membrane protein (HSP 16.3) from Mycobacterium tuberculosis (144 aa). SEEMS TO BELONG TO THE SMALL HEAT SHOCK PROTEIN (HSP20) FAMILY. (0)

31. PROBABLE LIPOPROTEIN LPQS Comments: Rv0847, (MTV043.40), length: 130 aa. Probable lpqS, lipoprotein. Contains possible signal sequence and PS00013 Prokaryotic membrane lipoprotein lipid attachment site. TBparse score is 0.920. (3)

32. Function: This protein is essential for replication of the chromosome. It is also involved in dna recombination and repair. Product: Probable single-strand binding protein ssb (helix-destabilizing protein) Evidence: experimental Proteomics: The putative product of this CDS corresponds to spot 3_210 identified in culture supernatant by proteomics at the Max-Planck-Institut fuer Infektionsbiologie, Germany (see citations below). Comments: Rv0054, (MTCY21D4.17), length: 164 aa. Probable ssb, single-strand binding protein, (2)

33. The electron transfer flavoprotein serves as a specific electron acceptor for other dehydrogenases. It transfers the electrons to the main respiratory chain via ETF-ubiquinone oxidoreductase (ETF dehydrogenase). Product: probable electron transfer flavoprotein (beta-subunit) FIXA (beta- ETF) (electron transfer flavoprotein small subunit) (ETFSS) Evidence:

experimental Proteomics: The putative product of this CDS corresponds to spot 3_122 identified in culture supernatant by proteomics at the Max-Planck-Institut fuer Infektionsbiologie (Germany), and spot FixA identified in short term culture filtrate by proteomics at the Statens Serum Institute (Denmark) (see citations below) (7)

34. Possibly a putative GTPase, modulating activity of HFLK and HFLC proteins. Product: Probable GTP-BINDING PROTEIN HFLX (EC 3.1.5.-) Comments: Rv2725c, (MTCY154.05c), length: 495 aa. Probable HFLX (HFL for high frequency of lysogenization), GTP-binding protein (7)

35. UNCTION Unknown Product: Conserved HYPOTHETICAL PROTEIN Evidence: experimental Proteomics: Identified by proteomics at the Max-Planck-Institut fuer Infektionsbiologie (Germany) (Functional category10 - conserved hypotheticals)

36. Function: Involved in dtdp-l-rhamnose biosynthesis [catalytic activity: dtdp-glucose = dtdp-4-dehydro-6-deoxy-d-glucose + h(2)o]. product: dtdp-glucose 4,6-dehydratase, RMLB (EC 4.2.1.46) Ma Y, Stern RJ, Scherman MS, Vissa VD, Yan W, Jones VC, Zhang F, Franzblau SG, Lewis WH, McNeil MR Drug targeting Mycobacterium tuberculosis cell wall synthesis: genetics of dTDP-rhamnose synthetic enzymes and development of a microtiter plate-based screen for inhibitors of conversion of dTDP-glucose to dTDP-rhamnose Antimicrob Agents Chemother (2001) 45:1407-16 (7)

37. Involved in purine salvage. Catalyses a salvage reaction resulting in the formation of AMP, that is energetically less costly than de novo synthesis [catalytic activity: amp + pyrophosphate = adenine + 5-phospho-alpha-d-ribose 1-diphosphate]. Product: ADENINE PHOSPHORIBOSYLTRANSFERASE APT (APRT) (AMP DIPHOSPHORYLASE) (AMP PYROPHOSPHORYLASE) (TRANSPHOSPHORIBOSIDASE) (7)

38. Catalyzes the reversible hydration of carbon dioxide [CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O]. Product: CARBONIC ANHYDRASE (CARBONATE DEHYDRATASE) (CARBONIC DEHYDRATASE) (7)

39. Possibly involved in transport of Na+/H+ across the membrane. Product: Probable conserved integral membrane transport protein YjcE Comments: Rv2287, (MTCY339.23c), length: 542 aa. Probable yjcE, conserved integral membrane transport protein, similar to eukaryote NA+/H+ exchangers (3)

40. Function unknown; probably involved in cellular metabolism. Product: possible hydrolase (7)

41. Function unknown; hydrolyses peptides. Product: Probable CYTOPLASMIC PEPTIDASE PEPQ (7)

42. Required for extrapulmonary dissemination. Mediates adherence to epithelial cells by binding to sulfated glycoconjugates present at the surface of these cells; binds heparin, dextran sulfate, fucoidan and chondroitin sulfate. Promotes hemagglutination of erythrocytes of certain host species. Induces mycobacterial aggregation. Product: HEPARIN BINDING HEMAGGLUTININ HBHA (ADHESIN) Evidence: experimental Comments: Rv0475, hbhA (MTCY20G9.01), length: 199 aa. hbhA, heparin-binding hemagglutinin (see citations below), equivalent to CAC31971.1|AL583925 possible hemagglutinin from Mycobacterium leprae (188 aa). Contains possible N-terminal signal sequence and K-A-rich region at C-terminus: SUBCELLULAR LOCATION: SURFACE ASSOCIATED. Pethe K, Alonso S, Biet F, Delogu G, Brennan MJ, Locht C, Menozzi FD The heparin-binding haemagglutinin of *M. tuberculosis* is required for extrapulmonary dissemination. Nature (2001) 412:190-194 Menozzi FD, Bischoff R, Fort E, Brennan MJ, Locht C Molecular characterization of the mycobacterial heparin-binding hemagglutinin, a mycobacterial adhesin Proc Natl Acad Sci USA (1998) 95:12625-12630 Menozzi FD, Rouse JH, Alavi M, Laude-Sharp M, Muller J, Bischoff R, Brennan MJ, Locht C Identification of a heparin-binding hemagglutinin present in Mycobacteria J Exp Med (1996) 184:993-1001 (3)

43. Not known. Elicits high level of INF-gamma from memory effector cells during the first phase of a protective immune response. Exported protein cotranscribed with Rv3874|MT3988|MTV027.09|LHP|CFP10. Product: 6 KDA EARLY SECRETORY ANTIGENIC TARGET ESAT6 (ESAT-6) Evidence: experimental Proteomics: The putative product of this CDS corresponds to spots 3875 identified in short term culture filtrate and cell wall by proteomics at the Statens Serum Institute (Denmark) (see citations from 2000 below). Comments: Rv3875, (MT3989, MTV027.10), length: 95 aa. esat6, Early Secretory Antigenic Target (see citations below) Gey Van Pittius NC, Gamieldien J, Hide W, Brown GD, Siezen RJ, Beyers AD The ESAT-6 gene cluster of Mycobacterium tuberculosis and other high G+C Gram-positive bacteria Genome Biol (2001) 2:44 Berthet FX, Rasmussen PB, Rosenkrands I, Andersen P, Gicquel BA Mycobacterium tuberculosis operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10)Microbiology (1998) 144:3195-203 Sørensen AL, Nagai S, Houen G, Andersen P, Andersen AB Purification and characterization of a low-molecular-mass T-cell antigen secreted by Mycobacterium tuberculosis. Infect Immun (1995) 63:1710-7 (3)

44. Secreted antigen 85-a FBPA (mycolyl transferase 85a) (fibronectin-binding protein a) (antigen 85 complex a) involved in cell wall mycoloylation. Proteins of the antigen 85 complex are responsible for the high affinity of mycobacteria to fibronectin. Possesses a mycolyltransferase activity required for the biogenesis of trehalose dimycolate (cord factor), a dominant structure necessary for maintaining cell wall integrity. Product: SECRETED ANTIGEN 85-A FBPA (MYCOLYL TRANSFERASE 85A) (FIBRONECTIN-BINDING PROTEIN A) (ANTIGEN 85 COMPLEX A) (EC 2.3.1.-) Evidence: experimental Proteomics: Corresponds to spots 3_494 and 3_496 identified in culture supernatant by proteomics at the Max-Planck-Institut fuer Infektionsbiologie (Germany), and spot 3804c identified in short term culture filtrate by proteomics at the Statens Serum Institute (Denmark) Puech V, Guilhot C, Perez E, Tropis M, Armitige LY, Gicquel B, Daffe M Evidence for a partial redundancy of the fibronectin-binding proteins for the transfer of mycoloyl residues onto the cell wall arabinogalactan termini of Mycobacterium tuberculosis. Mol Microbiol (2002) 44:1109-22 Harth G, Lee BY, Wang J, Clemens DL, Horwitz MA Novel insights into the genetics, biochemistry, and immunocytochemistry of the 30-kilodalton major extracellular protein of Mycobacterium tuberculosis Infect Immun (1996) 64:3038-47 Borremans M, de Wit L, Volckaert G, Ooms J, de Bruyn J, Huygen K, van Vooren JP, Stelandre M, Verhofstadt R, Content J Cloning, sequence determination, and expression of a 32-kilodalton-protein gene of Mycobacterium tuberculosis Infect Immun (1989) 57:3123-30 (1)

45. Cell surface lipoprotein mpt83 precursor (lipoprotein p23) Function: not really known. Product: CELL SURFACE LIPOPROTEIN MPT83 PRECURSOR (LIPOPROTEIN P23) Evidence: experimental Comments: Rv2873, (MTCY274.04), length: 220 aa. mpt83 (alternate gene name: mpb83), cell surface lipoprotein (see citations below). Juarez MD, Torres A, Espitia C Characterization of the Mycobacterium tuberculosis region containing the mpt83 and mpt70 genes(1)FEMS Microbiol Lett (2001) 203:95-102Vosloo W, Tippoo P, Hughes JE, Harriman N, Emms M, Beatty DW, Zappe H, Steyn LM. Characterization of a lipoprotein in Mycobacterium bovis (BCG) with sequence similarity to the secreted protein MPB70Gene (1997) 188:123-8Matsuo T, Matsuo H, Ohara N, Matsumoto S, Kitaura H, Mizuno A, Yamada T Cloning and sequencing of an MPB70 homologue corresponding to MPB83 fr Mycobacterium bovis BCG. Scand J Immunol (1996) 43:483-9[MEDLINE: N/A] [RELATED GENES]Hewinson RG, Michell SL, Russell WP, McAdam RA, Jacobs Jr. WR. Molecular characterization of MPT83: a seroreactive antigen of Mycobacterium tuberculosis with homology to MPT70Scand J Immunol (1996) 43:490-9 (3)

46. Function: not really known. Product: MAJOR SECRETED IMMUNOGENIC PROTEIN MPT70 PRECURSOR Comments: Rv2875, (MTCY274.06), length: 193 aa. mpt70 (alternate gene name: mpb70), major secreted immunogenic protein MPT70 precursor (see citations below) Juarez MD, Torres A, Espitia C. Characterization of the Mycobacterium tuberculosis region containing the mpt83 and mpt70 genes(1)FEMS Microbiol Lett (2001) 203:95-102Matsuo T, Matsumoto S, Ohara N, Kitaura H, Mizuno A, Yamada TDifferential transcription of the MPB70 genes in two major groups of Mycobacterium bovis BCG substrains. Microbiology (1995) 141 (Pt 7):1601-7Matsumoto S, Matsuo T, Ohara N, Hotokezaka H, Naito M, Minami J, Yamada T. Cloning and sequencing of a unique antigen MPT70 from Mycobacterium tuberculosis H37Rv and expression in BCG using Escherichia coli-mycobacteria shuttle vector. Scand J Immunol (1995) 41:281-7Radford AJ, Wood PR, Billman-Jacobe H, Geysen HM, Mason TJ, Tribbick G. Epitope mapping of the Mycobacterium bovis secretory protein MPB70 using overlapping peptide analysis. Gen Microbiol (1990) 136:265-72Radford AJ, Duffield BJ, Plackett PCloning of a species-specific antigen of Mycobacterium bovis. Infect Immun (1988) 56:921-5 (3)

47. Function: binds to cpn60 in the presence of mg-ATP and suppresses the ATPase activity of the latter. Product10 KDA CHAPERONIN GROES (PROTEIN CPN10) (PROTEIN GROES) (BCG-A HEAT SHOCK PROTEIN) (10 KDA ANTIGEN) Evidence: experimental Proteomics: The product of this CDS corresponds to spots 5_154, 5_160, 5_157, 5_159 and 5_152 identified in culture supernatant by proteomics at the Max-Planck-Institut fuer Infektionsbiologie (see citations from 1999 below), and spots 3418c identified in short term culture filtrate, cell wall and cytosol by proteomics at the Statens Serum Institute (Denmark) (see citations from 2000 below). Kong TH, Coates AR, Butcher PD, Hickman CJ, Shinnick TM Mycobacterium tuberculosis expresses two chaperonin-60 homologs Proc Natl Acad Sci U S A (1993) 90:2608-12 Baird PN, Hall LM, Coates AR A major antigen from Mycobacterium tuberculosis which is homologous to the heat shock proteins groES from Escherichia coli and the htpA gene product of Coxiella burneti. Nucleic Acids Res (1988) 16:9047 (0)

48. Function: unknown Product: PPE FAMILY PROTEIN Evidence: experimental Comments: Rv1196, (MTCI364.08), length: 391 aa. PPE18 (alternate gene name: mtb39a). Member of the Mycobacterium tuberculosis PPE family of glycine-rich proteins, highly similar to others Dillon DC, Alderson MR, Day CH, Lewinsohn DM, Coler R, Bement T, Campos-Neto A, Skeiky YA, Orme IM, Roberts A, et al. Molecular characterization and human T-cell responses to a member

of a novel Mycobacterium tuberculosis mtb39 gene family Infect Immun (1999) 67:2941-50 (Functional category 6 - PE/PPE)

49. Function Unknown Product: PE FAMILY PROTEIN (6)

50. Function Unknown Product: PE FAMILY PROTEIN (6)

51. Function Unknown Product: PE FAMILY PROTEIN (6)

52. Function Unknown Product: PE-PGRS FAMILY PROTEIN Comments: Rv3508, (MTV023.15), length: 1901 aa. Member of the Mycobacterium tuberculosis PE family, PGRS subfamily of Gly-rich proteins, similar to others from *M. tuberculosis* strains H37Rv and CDC1551 (6)

53. Function Unknown Product: PE FAMILY PROTEIN (6)

54. Function Unknown Product: Low molecular weight protein antigen 7 cfp7 (10 kDa antigen) (CFP-7) (Protein TB10.4) Evidence: experimental Proteomics: The putative product of this CDS corresponds to spots 0288 identified in short term culture filtrate by proteomics at the Statens Serum Institute (Denmark) Skjot RL, Oettinger T, Rosenkrands I, Ravn P, Brock I, Jacobsen S, Andersen P Comparative evaluation of low-molecular-mass proteins from Mycobacterium tuberculosis identifies members of the ESAT-6 family as immunodominant T-cell antigens Infect Immun (2000) 68:214-220 Rindi L, Lari N, Garzelli C Search for genes potentially involved in Mycobacterium tuberculosis virulence by mRNA differential display Biochem Biophys Res Commun (1999) 258:94-101 (Functional category10 - conserved hypotheticals)

55. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10)

56. Function: Unknown Product: HYPOTHETICAL PROTEIN Comments: Rv3889c, (MTCY15F10.23), length: 276 aa. Hypothetical unknown protein. (Functional category 8 - unknown)

57. Function Unknown Product: Conserved HYPOTHETICAL PROLINE AND ALANINE RICH PROTEIN (10)

58. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN Comments: Rv3182, (MTV014.26), length: 114 aa. Hypothetical protein, with some similarity to other hypothetical bacterial proteins (10)

59. Function: Unknown Product: HYPOTHETICAL PROTEIN (8)

60. Function: Unknown Product: HYPOTHETICAL PROTEIN (8)

61. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN Evidence: experimental Proteomics: Identified by proteomics (see citation below). (10)

62. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10)

63. Eis protein Function Unknown Product: Conserved HYPOTHETICAL PROTEIN Evidence: experimental Proteomics: Identified by proteomics at the Statens Serum Institute (Denmark)
64. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10)
65. Function: Unknown Product: Probable phiRv2 PROPHAGE PROTEIN Comments: Rv2657c, (MTCY441.26c), length: 86 aa. Probable phiRv2 phage protein (excisionase) (see citation below), Hatfull GH, Jacobs Jr. WR Molecular Genetics of Mycobacteria; Second chapter: Molecular Genetics of Mycobacteriophages ASM Press, ISBN 1-55581-191-4 (2000) :37-54 (Functional category 5 - insertion seqs and phages)
66. Function Unknown Product: CADMIUM INDUCIBLE PROTEIN CADI Evidence: experimental Proteomics: Identified by proteomics Hotter GS, Wilson T, Collins DM Identification of a cadmium-induced gene in Mycobacterium bovis and Mycobacterium tuberculosis FEMS Microbiol Lett (2001) 200:151-5 (10)
67. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10)
68. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10)
69. Function: Unknown Product: HYPOTHETICAL PROTEIN Comments: Rv1955, (MTCY09F9.09c), length: 170 aa. Hypothetical unknown protein, start overlaps another ORF, Rv1954c (8)
70. Function: Thought to promote the resuscitation and growth of dormant, nongrowing cell. Could also stimulates the growth of several other high G+C gram+ organisms, *e.g.* Mycobacterium avium, Mycobacterium bovis (BCG), Mycobacterium kansasii, Mycobacterium smegmatis. Product: Probable RESUSCITATION-PROMOTING FACTOR RPFC Comments: Rv1884c, (MTCY180.34), length: 176 aa. Probable rpfC, resuscitation promoting factor (see citation below) Mukamolova GV, Kaprelyants AS, Young DI, Young M, Kell DB A bacterial cytokine. Proc Natl Acad Sci U S A (1998) 95:8916-21 (3)
71. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10)
72. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10)
73. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10)
74. Function: Unknown Product: HYPOTHETICAL PROTEIN (8)
75. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10)
76. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10)
77. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10)
78. Function: Involved in TRANSCRIPTIONAL MECHANISM. Product: POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN WHIB-LIKE WHIB6 Comments: Rv3862c,

(MTCY01A6.06), length: 116 aa. Possible whiB6 (alternate gene name: whmF), WhiB-like regulatory protein (see citation below), similar to WhiB paralogue of Streptomyces coelicolor, wblE gene product (85 aa) Hutter B, Dick T Molecular genetic characterization of whiB3, a mycobacterial homologue of a Streptomyces sporulation factor Res Microbiol (1999) 150:295-301 (Functional category 9 - regulatory proteins)

79. Function Unknown Product: Conserved HYPOTHETICAL PROTEIN (10).

References:

1. World Health Organization (1998) The WHO/IUATLD Global Project on Anti-tuberculosis Drug Resistance Surveilance 1994-1997

2. Raviglione, M. C. (2002) 4th World Congress on Tuberculosis Program/WHO, 2.

3. Hu, Y. M., Butcher, P. D., Sole, K., Mitchison, D. A. and Coates, A. R. (1998) FEMS Microbiol Lett 158, 139.

4. Primm, T. P., Andersen, S. J., Mizrahi, V., Avarbock, D., Rubin, H. and Barry, C. E. (2000) J Bacteriol 182, 4889.

5. Avarbock, D., Salem, J., Li, L. S., Wang, Z. M. and Rubin, H. (1999) Gene 233, 261.

6. Avarbock, D., Avarbock, A. and Rubin, H. (2000) Biochemistry 39 11640.

8. O'Brien, W. (1976) Anal. Biochem. 76, 423.

10. Gentry, D. R., and Cashel, M. (1996) Molecular Microbiology 19, 1373.

11. Mittenhuber, G. (2001) J. Mol. Microbiol. Biotechnol. 3, 585.

12. Mechold, U., Murphy, H., Brown, L., and Cashel, M. (2002) Journal of Bacteriology 184, 2878.

13. Aravind, L., and Koonin, E. V. (1998) Trends Biochem. Sci. 23, 469.

14. Rost, B., and Sander, C. (1994) Proteins 19, 55.

14.2 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res 25, 3389.

14.4 Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998) Proc. Natl. Acad. Sci. USA 95, 5857.

14.6 Letunic, I., Goodstadt, L., Dickens, N. J., Doerks, T., Schultz, J., Mott, R., Ciccarelli, F., Copley, R. R., Ponting, C. P., and Bork, P. (2002) Nucleic Acids Res. 30, 242.

15. CGMP – Turko, I. V., Francis, S. H., and Corbin, J. D. (1998) Journal of Biological Chemistry 273, 6460.

16. Gropp, M., Strausz, Y., Gross, M., and Glaser, G. (2001) Journal of Bacteriology 183, 570.

17. Wendrich, T. M., and Marahiel, M. A. (1997) Mol. Microbiol. 26, 65.

18. Metzger, S., Schreiber, G., Aizenman, E., Cashel, M., and Glaser, G. (1989) J. Biol. Chem. 264, 21146.

19. Schreiber, G., Metzger, S., Aizenman, E., Roza, S., Cashel, M., and Glaser, G. (1991) Journal of Biological Chemistry 266, 3760.

20. Yang, X., and Ishiguro, E. E. (2001) Biochem. Cell. Biol. 79, 729.

21. Wolf, Y. I., Aravind, L., Grishin, N. V., and Koonin, E., V. (1999) Genome Res. 9, 689.

22. Aravind, L., and Koonin, E. V. (1999) J. Mol. Biol 287, 1023.

23. Wek, R. C., Jackson, B. M., and Hinnebusch, A. G. (1989) Proc. Natl Acad. Sci. USA 86, 4579.

24. Sood, R., Porter, A. C., Olsen, D., Cavener, D. R., and Wek, R. C. (2000) Genetics 154, 787.

25. Wek, S. A., Zhu, S., and Wek, R. C. (1995) Mol. Cell. Biol. 15, 4497.

26. Dong, J., Qiu, H., Garcia-Barrio, M., Anderson, J., and Hinnebusch, A. G. (2000) Molecular Cell 6, 269

27. Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C., and Moras, D. (1991) Science 252, 1682.

28. Hinnebusch, A. (1994) Semin. Cell Biol. 5, 417.

29. Hinnebusch, A. (1996) In Translational Control (Hershey, M., Mathews, M., and Sonenberg, N., Eds.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 199.

30. Cashel, M., Gentry, D. R., Hernandez, V. J., and Vinella, D. (1996) In Escherichia coli and Salmonella: Cellular and Molecular Biology (Neidhardt, F. C., Curtis, R., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., Eds) ASM Press, Washington, DC.

31. Qiu, H., Dong, J., Hu, C., Francklyn, C. S., and Hinnebusch, A. G. (2001) EMBO Journal 20, 1425.

32. Yang, X., and Ishiguro, E. (2001) Journal of Bacteriology 183, 6532.

33. Eriani, G., Cavarelli, J., Martin, F., Dirheimer, G., Moras, D., and Gangloff, J. (1993) Proc. Natl. Acad. Sci. USA 90, 10816.

34. Friesen, J. D., Fiil, N. P., Parker, J. M., and Haseltine, W. A. (1974) Proc. Natl. Acad. Sci. USA 71, 3465.

35. Parker, J., Watson, R. J., and Friesen, J. D. (1976) Mol. Gen. Genet. 144, 111.